NASA/TM–2003-212269

Detailed Vibration Analysis of Pinion Gear with Time-Frequency Methods

Marianne Mosher
Ames Research Center, Moffett Field, California

Anna H. Pryor
Ames Research Center, Moffett Field, California

David G. Lewicki
U.S. Army Research Laboratory, NASA Glenn Research Center, Cleveland, Ohio

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

June 2003

DETAILED VIBRATION ANALYSIS OF PINION GEAR WITH TIME-FREQUENCY METHODS

Marianne Mosher, Anna H. Pryor and David G. Lewicki*

NASA Ames Research Center

SUMMARY

In this paper, the authors show a detailed analysis of the vibration signal from the destructive testing of a spiral bevel gear and pinion pair containing seeded faults. The vibration signal is analyzed in the time domain, frequency domain and with four time-frequency transforms: the Short Time Frequency Transform (STFT), the Wigner-Ville Distribution with the Choi-Williams kernel (WV-CW), the Continuous Wavelet Transform (CWT) and the Discrete Wavelet Transform (DWT). Vibration data of bevel gear tooth fatigue cracks, under a variety of operating load levels and damage conditions, are analyzed using these methods. A new metric for automatic anomaly detection is developed and can be produced from any systematic numerical representation of the vibration signals. This new metric reveals indications of gear damage with all of the time-frequency transforms, as well as time and frequency representations, on this data set. Analysis with the CWT detects changes in the signal at low torque levels not found with the other transforms. The WV-CW and CWT use considerably more resources than the STFT and the DWT. More testing of the new metric is needed to determine its value for automatic anomaly detection and to develop fault detection methods for the metric.

Introduction

Many metrics based on frequency analysis are currently used on vibration data to detect faults from gearboxes [1]. Although these methods have been shown to find faults [2], traditional methods such as spectrum analysis, waterfall plot, cepstrum analysis, and matched filtering [3], were developed for use on stationary data. A signal has weak stationarity when its mean and variance are constant, and strong stationarity when all its moments are constant. Stationarity used by itself is generally understood to mean weak stationarity. However, many systems are not stationary. For instance, in Health and Usage Monitoring Systems (HUMS)[4], the vibrational frequencies being measured by the accelerometers can change rapidly in time, especially if a fault has occurred. Many types of gear damage produce localized changes in the signal so that the signal is no longer stationary on the time-scale of the gear tooth meshing. The signal near the meshing with a defective tooth may vary

* U.S. Army Research Laboratory NASA Glenn Research Center Cleveland, OH

considerably from the rest of the signal. If the situation is critical, it is important to determine the severity so that corrective actions can be taken.

In standard Fourier analysis, for example, a signal is decomposed into individual frequencies. Unfortunately, there is no way to determine when each of those frequencies has occurred. However, there are signal processing methods that give local information about both time and frequency. These methods localize signal features in both time and frequency; therefore, these methods have the potential to be more sensitive to early changes in the signal due to impending faults.

Many time-frequency (TF) methods have been applied to the detection of faults in gears. Wang has claimed the Short Time Fourier Transform (STFT) to be a powerful tool in detecting local gear damage at an early state [5]. In periodic data, strong harmonics may obscure small transient events. According to one study, STFTs will perform better here than conventional methods [6]. Others have investigated STFT methods for the early detection of faults in gears [7,8]. Local tooth faults in spur gears have been found with the Wigner-Ville (WV-CW) method [1, 9]. Other researchers have been inspired by WV-CW and have modified it to create their own methods [10,11,6]. The Continuous Wavelet Transform (CWT) is being studied more and its usefulness increases. In one study, the phase map of the wavelet transform was found to have distinctive features near a cracked tooth [12]. In another study, satisfactory results were obtained from a de-noising method based on wavelet analysis used for the diagnosis of mechanical vibration signals [13]. Many others use the CWT as applied to the detection of faults in gears [14, 5, 15, 16, 17, 18, 19, 20, 21, 22, 23]. The Discrete Wavelet Transform (DWT) or Multi Resolution Analysis has also had some good results with gear faults. In one study, when the data was preprocessed using the DWT and then fed into a multi-layer neural network, gear faults were successfully found and classified into different groups [24]. In another study, the DWT was used to create a residual error. The probability density function of the residual error was then expanded into a Hermite polynomial and the coefficients were used as a feature vector to estimate early fatigue cracks in gears [3]. Other researchers have had both mixed and good results using the DWT in gear fault detection [25,4].

It is extremely difficult to use a TF method as it stands alone for HUMS analysis. For one thing, the dimensionality is simply too large. In addition, there are no obvious fault detection criteria, other than human observation, without some dimensionality reduction. Some authors have reduced the dimensionality without creating a metric [13,8]. Other researchers have reduced the input into a neural net and used it as a classifier [24,1, 15, 16, 17]. Because of the way the nets are trained however, the nets may not find faults that differ from those in the training set since the nets are trained on only a few examples of faults. On the other hand, some authors have both reduced the dimensionality and created metrics [14,3,26,4,6,7]. For example, in [6], Williams and Zalubas use a generalized and improved variation of Choi-Williams called RID. They analyzed a spectrogram with singular value decomposition. Then they looked at scatter plots of the first two principal components. Williams also examined a metric produced by projecting onto what Williams calls the "zero-subspace" and what the current authors interpret as being the noise subspace for the normal and faulted data. They also studied a squared distance measure between the projection of faulty and normal signals. In another method, Ferlez [19], selected one scale from the CWT and used the statistical characteristics of the probability density function of the selected scale to separate fault from no fault condition.

The objectives of this report are to examine the four time-frequency transforms of gear vibration signals with and without damage and to develop a metric based on the time-frequency transforms that will detect anomalies in the signals. This report gives more details than the first study by the authors in [27]. The four time-frequency methods are compared here by first analyzing them on a set of experimental data which will be called the Multiple-Seeded Fault Data. Many examples of vibration are shown for the four Time-Frequency methods as well as time history and Fourier transform in order to examine the changes in the vibration as the test progressed. Features in these various representations will be described and compared. A simple measure of the changes in these representation will be developed for anomaly detection by means of modeling with singular value decomposition (SVD). anomoly detection occurs through deviations of measurements from the model. Finally, all four methods are compared with each other and with time and frequency analyses using a simple one-dimensional metric developed by the authors. The following are included: the assumptions, resources, resolution, speed of calculation and results of analyses.

DESCRIPTION OF TEST AND ANALYSIS

Test

The vibration data reported in the current work were acquired from tests of an OH-58A helicopter main-rotor transmission (fig. 1) in the NASA Glenn Research Center500-hp Helicopter Transmission Test Facility (fig. 2). The objective of the tests was to supply experimental data to study the detection of gear tooth fatigue cracks and their propagation rates under a variety of operating load levels. The gear under study was the spiral bevel pinion with 19 teeth (fig. 3a). Nine notches were fabricated in the fillets on various teeth of the pinion. The notch sizes varied from 0.5 to 0.1" in length, and 0.015 to 0.045" in depth. The pinion was run 4.9 million cycles at 6060 rpm and torque levels from 2479 to 4649 in-lb (80, 100, 125 and 150% design load), see table 1. By the end of the test, after 1.9 million cycles at maximum torque, five teeth fractured off the pinion (fig 3b). More information about the transmission and test rig can be found in [28].

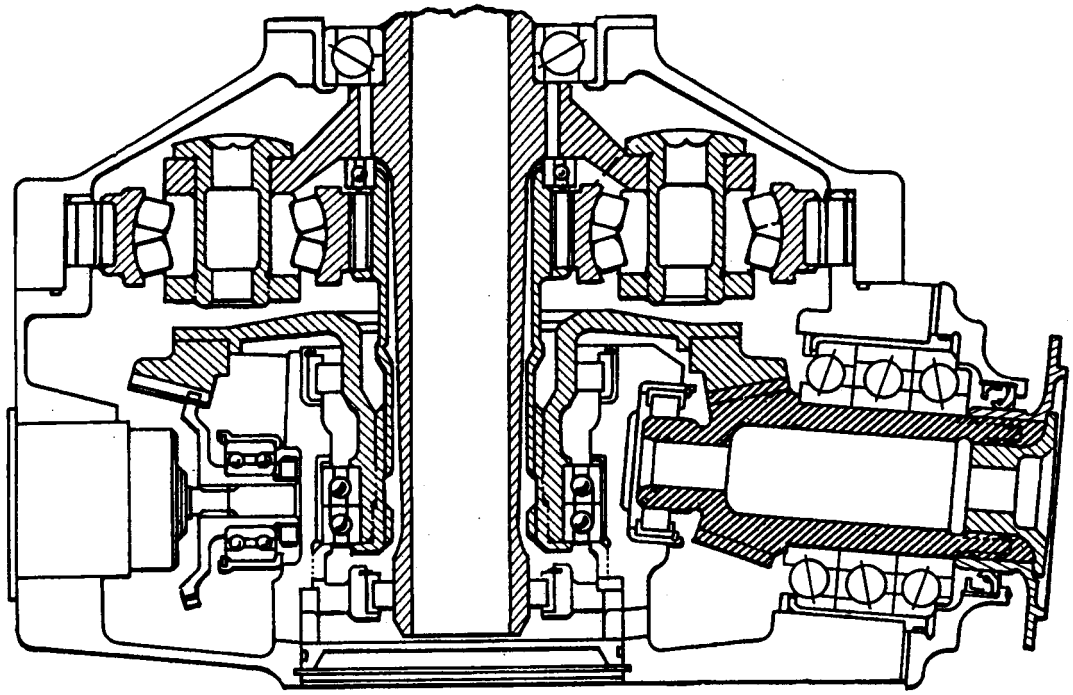


Figure 1. OH-58A Helicopter Main-Rotor Transmission

Table 1. OH-58A Pinion Operating Conditions

Data Record	1-43	44-126	127-209	210-449
Time (cycles/million)	1	1	1	1.9
Speed (rpm)	6060	6060	6060	6060
Torque (in-lb)	2479	3099	3874	4649
Torque (%)	80	100	125	150

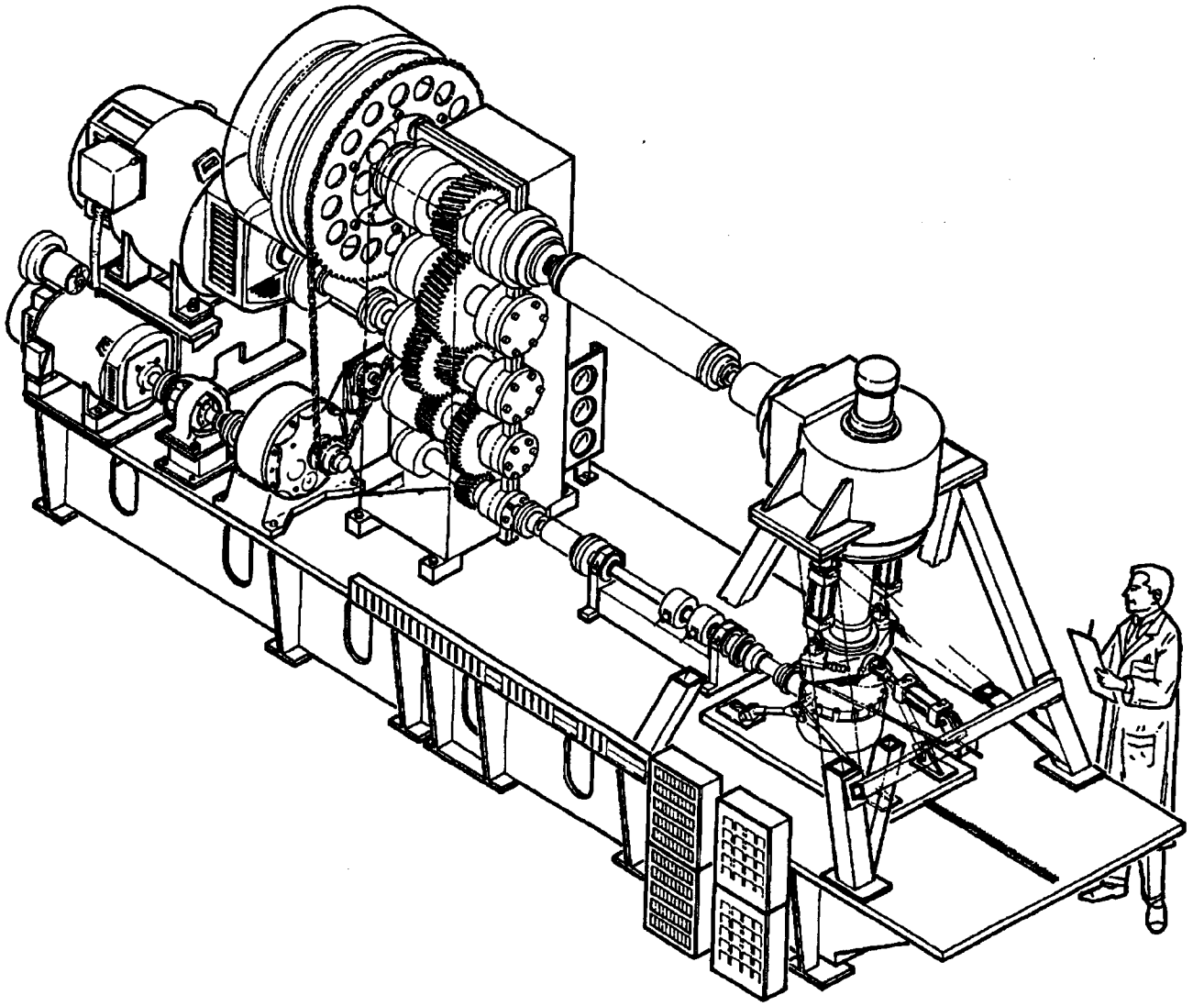


Figure 2. NASA Glenn 500-hp Helicopter Transmission Test Facility

During the tests, vibration data were periodically collected from the transmission. Five piezoelectric accelerometers were mounted on the OH-58A transmission housing (fig. 4). The accelerometers were lightweight, high frequency, and had integral electronics. They had a nominal resonant frequency of 90 kHz and an output sensitivity of 10 mV/g. Measurements from accelerometer number 1 are used in the current study. Accelerometer number 1 is oriented radial to the input shaft and is close to the pinion. The synchronously averaged signal from accelerometer number 1 showed consistently higher amplitude than signals from the other accelerometers. Because the pinion contained 19 teeth, the vibration signals from the healthy gear are expected to contain a waveform with 19 repetitions per rotation and significant frequency components at 19 times the shaft rotation and its harmonics.

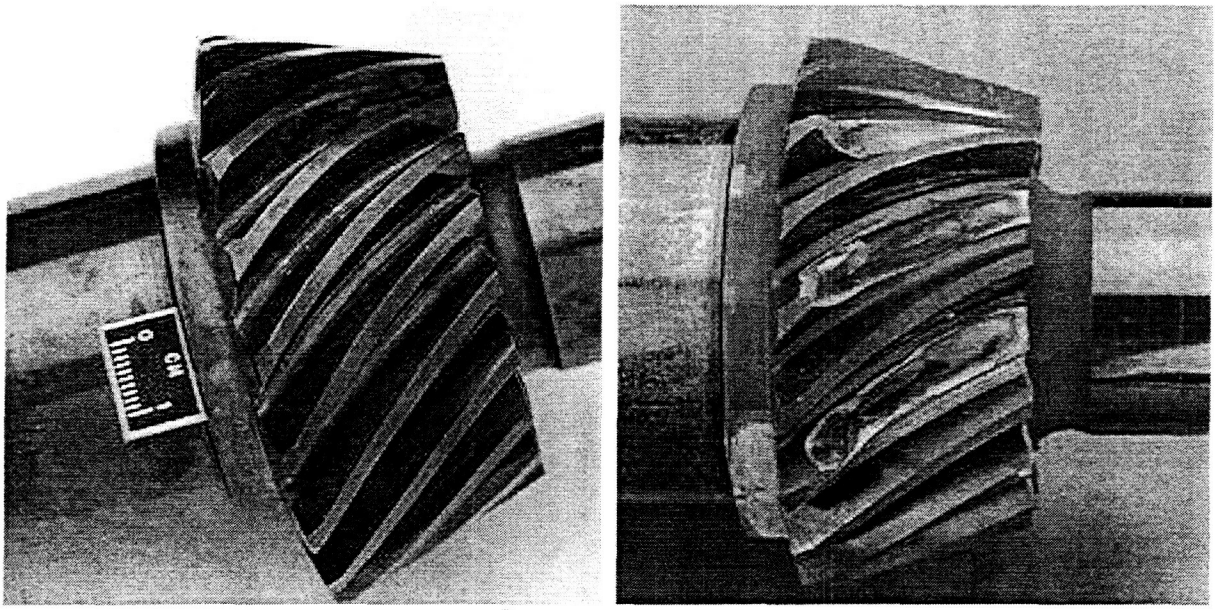


Figure 3. OH-58A Spiral-Bevel Pinion. (a) Normal teeth (b) fractured teeth

The vibration data were recorded on magnetic tape using high-frequency wideband group II electronics approximately every 180,000 cycles (every 30 minutes). The data were then post processed. Data were interpolated to1024 points per rotation of the pinion and then synchronously averaged. By synchronously averaging, the components of the signal not synchronous with the pinion rotation were reduced. This procedure reduces both the noise and the signals from other hardware components. Over the short span of one record the vibration signal is relatively stationary on the time scale of the gear rotation because the test rig is operated under very tight conditions, so averaging on the time-scale of the gear rotation is reasonable. Local signal components due to gear damage will be preserved as long as the signal repeats each rotation of the gear. In general, the fault damage is not expected to change much over the few seconds of a data record.

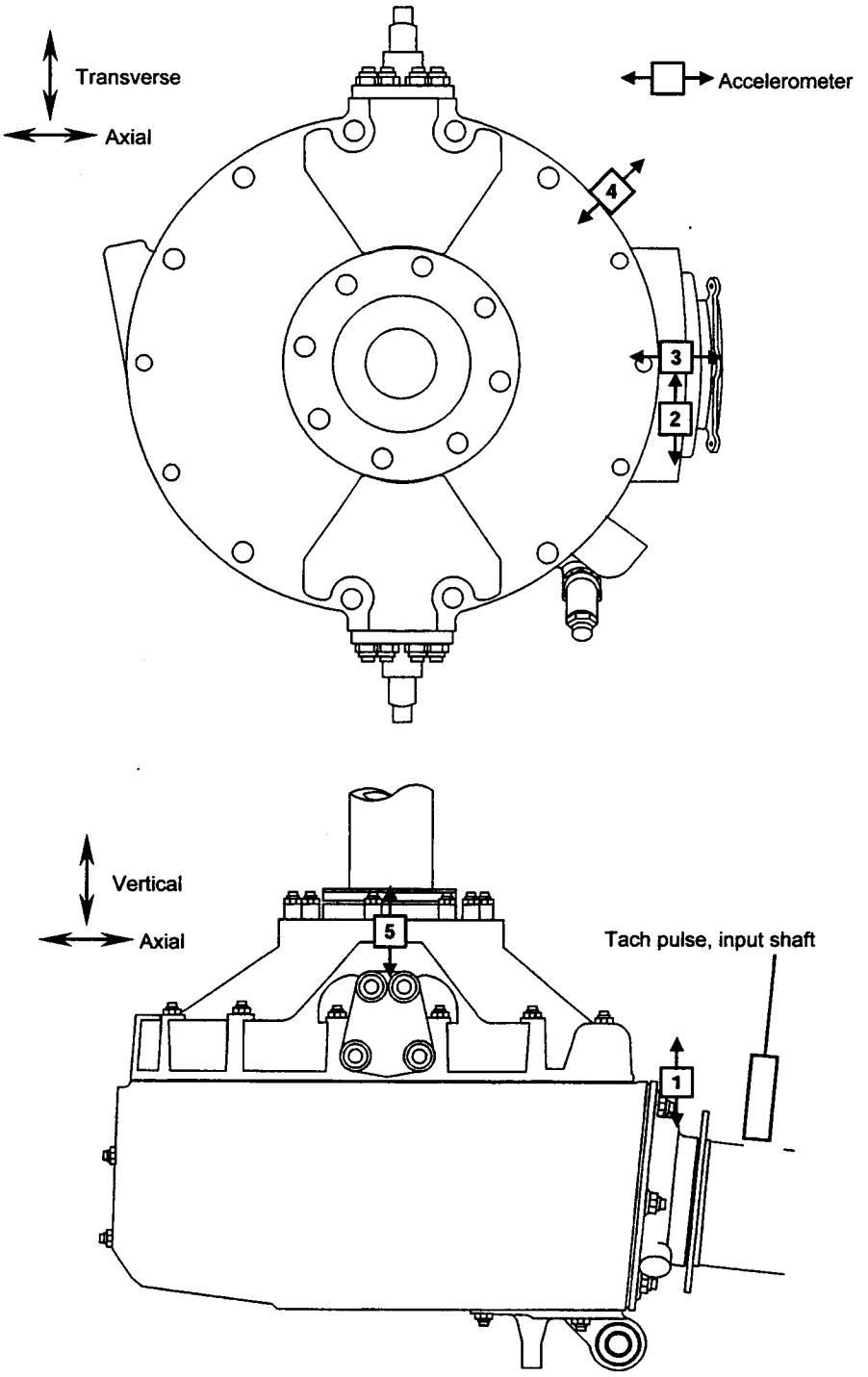


Figure 4. Accelerometer Locations on OH-58A Transmission.

Overview of Analysis

The current study began with the synchronously averaged time histories. The data were then low-pass filtered and decimated by an order of four to 256 points per pinion rotation for the analyses in this paper. It was not practical to perform analyses on the longer blocksizes with some of the TF methods. In addition, the higher frequencies did not appear to be needed to find changes in the data due to gear damage.

In this study, the term time-frequency will be used. The presentations of the data are actually sampled at uniform phase spacing of the gear instead of uniform time. This keeps the signal in phase with the gear teeth even though the rpm may change. In addition, the W-V distribution and the Wavelet transforms do not transform directly into frequency, but something that can be related to frequency.

The analysis in this study projects the synchronously averaged time history onto a time-frequency space. All time-frequency transforms in this study except the orthogonal wavelet transform, project the vibration signal onto a space of much higher dimension and thus produce a very redundant representation of the signal. The elements in the time-frequency space are signals with localized frequency and time content. Local features in the vibration signal can be represented in the time-frequency space. Projecting the time-frequency representation of the signal onto a very low-dimensional subspace produces a model. This low-dimensional subspace of the very large time-frequency space is optimized to represent the signal in a training set of the data. A metric that identifies anomalies is produced based upon how closely the projection of the transformed signal onto the very small subspace matches the original signal. This metric is a measure of the distance of a signal from the model.

All of the computational analyses were done with the mathematical software package MATLAB.

Time-Frequency Methods

Formulas for continuous time transforms offer a more concise notation than those of discrete time. Although this paper contains continuous time formulas, the discrete time analogs were used for the computation. As far as a bit of notation, if $f(x)$ is a function, then $\bar{f}(x)$ is its complex conjugate.

Short Time Fourier Transform (STFT)

The Short Time Fourier Transform (STFT), also known as the Windowed Fourier Transform or spectrogram, is a development that extends standard Fourier Transform techniques to handle non-stationary data [29]. Fourier Transforms are applied to short windows of data. These windows are moved along the data and may overlap. This transforms one-dimensional data into two-dimensional data, one dimension for frequency and a second dimension for window location in the data. The STFT gives information for a fixed frequency band and time resolution dependent on the window. An impulsive event appears in an STFT as increased levels for all frequencies at the time of the

impulse. The limitation of the STFT is that signal components having a poor match to the fixed time and frequency resolution will be obscured.

Given a time interval $T > 0$, let $g(u)$ be a function that vanishes outside the interval $-T \le u \le 0$. The STFT of a function $f(u)$ is defined to be

$$\tilde{f}(\omega,t) = \int_{-\infty}^{\infty} \bar{g}(u-t) f(u) e^{-2\pi i \omega u} du \tag{1}$$

The analyses in this paper use a Hanning window of 32 points in length with 24 points of overlap for the function $g(u)$.

Choi-Williams Distribution

The Wigner-Ville Distribution (WV) is a nonlinear transform that maps one-dimensional data into two dimensions. One dimension has frequency-like characteristics and the other dimension has time-like characteristics. Integrating the WV along the frequency-like dimension produces the squared amplitude of the original signal. Integrating the WV along the time-like dimension produces the squared amplitude of the Fourier Transform of the original signal. The WV has units of signal power, not amplitude. Phase information is no longer available.

As one of the first time frequency methods to be used, the WV distribution has been widely studied. Given a function $f(u)$, the Wigner-Ville Distribution is defined as

$$W(t,\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{f}\left(t - \frac{u}{2}\right) f\left(t + \frac{u}{2}\right) e^{-i\omega u} du \tag{2}$$

Unfortunately, if the signal is even moderately complicated, there will be significant cross-terms, which will make interpretation nearly impossible. The Choi-Williams Distribution (WV-CW) is a modification of the Wigner-Ville Distribution [30,31]. The Choi-Williams Distribution uses a kernel (an additional multiplied term inside equation 2) that dampens out the cross-terms. While still present, the cross-terms are significantly dampened and cause lesser problems in interpreting the signal. The kernel that Choi and Williams chose was

$$\phi(\theta,\tau) = e^{-\phi^2 \tau^2 / \sigma^2} \tag{3}$$

where the parameter σ controls the amount of attenuation. The amplitude of the cross-terms is directly proportional to σ. However, if the cross-terms are suppressed too much, then the auto-terms will lose resolution in the time frequency plane. The final Choi-Williams Distribution is defined as

$$\begin{aligned} W(t,\omega) &= \frac{1}{2\pi} \int_{\tau} e^{-i\tau\omega} K(t,\tau;\sigma) d\tau \\ K(t,\tau;\sigma) &= \int_{\xi} \int_{\mu} e^{i\xi(t-\mu)} \bar{f}\left(\mu - \frac{\tau}{2}\right) f\left(\mu + \frac{\tau}{2}\right) e^{-\xi^2 \tau^2 / \sigma} d\mu d\xi \end{aligned} \tag{4}$$

For this study $\sigma = 0.05$.

Continuous Wavelet Transform

The Continuous Wavelet Transform (CWT) is a time-frequency method that builds on the idea of the STFT. Whereas the STFT is limited in resolution because of its fixed window size, the CWT uses a variable window size with short windows for high frequencies and long windows for low frequencies. This allows the CWT to be a powerful tool in representing local features of a signal that other methods such as the STFT may miss entirely.

Given a mother wavelet, all other wavelets are made by dilating and/or translating the mother wavelet. These wavelets are moved along the data and overlap. This transforms the data into two dimensions. The first is for scale (the size of the wavelet) and the second is for the location in the data. For each wavelet, the scales can be converted back to a range of frequencies dominate the spectrum in the scale.

In continuous wavelet analysis [29, 32], begin with a complex-valued window function $\Psi(t)$ called the mother wavelet. Given an arbitrary $p \geq 0$ and any real number $s \neq 0$, the scaled wavelet is defined as

$$\Psi_s(u) \equiv |s|^{-p} \Psi\left(\frac{u}{s}\right) \tag{5}$$

In order to describe a given signal at a local time, the signal is compared to translated versions of Ψ_s. Given an interval of length T near u=0, if $\Psi(u)$ is supported (non-zero) on T, then near $u = 0$, if $\Psi_s(u)$ is supported on an interval of length $|s|T$ then the function

$$\Psi_{s,t}(u) \equiv \Psi_s(u-t) = |s|^{-p} \Psi\left(\frac{u-t}{s}\right) \tag{6}$$

is supported on $|s|T$ near $u = t$. Assuming that the mother wavelet belongs to $L^2(\mathbf{R})$ then $\Psi_{s,t}$ does also. If a given function $f(t)$ also belongs to $L^2(\mathbf{R})$, then the continuous wavelet transform (CWT) is defined as

$$\tilde{f}(s,t) \equiv \int_{-\infty}^{\infty} \overline{\Psi}_{s,t}(u) f(u) du \tag{7}$$

For this study the seventh order Daubechies wavelet scaling filter, known as db7 was used on all of the data records.

Discrete Wavelet Transform

Given functions $f(x)$ and $\Psi(x)$ such that $\int_{-\infty}^{\infty} \Psi(t) dt = 0$, the discrete wavelet transform of $f(x)$ is defined to be

$$\tilde{f}(a,b) = a^{-m/2}\int_{-\infty}^{\infty} \Psi\left(a^{-m}t - nb\right) f(t)dt \tag{8}$$

It is then divided into two cases. In the first case, called redundant discrete systems or frames, the dilation parameter, a, and the translation parameter, b, are discrete. In particular, for a, only powers of one fixed dilation parameter greater than one are used. This discrete system is not used in this paper and will not be discussed further. In the second case, called orthogonal wavelet, orthonormal wavelet or multi-resolution analysis, specific choices of a, b, and Ψ are made so that the wavelets that are created form an orthogonal or orthonormal basis. For example, if $a = 2$ and $b = 1$, then there exists Ψ such that

$$\Psi_{m,n}(x) = 2^{-m/2}\Psi\left(2^{-m}x - n\right) \tag{9}$$

form an orthonormal basis for $L^2(\mathbf{R})$.

The Multi-Resolution Analysis is a fast iterative algorithm that analyzes any signal. At each level of decomposition, a signal is high-pass filtered and down-sampled to produce one of the detail signals comprising the wavelet analysis. The same signal is also low-pass filtered and down-sampled to produce the next averaged signal to be used as part of the wavelet analysis. This whole process is repeated on the average signal enough times for the necessary analysis of the signal.

In this study, the seventh order Daubechies wavelet scaling filter, known as db7, was used on the discrete wavelet transform. There were seven levels evaluated. This produces seven details and one average. The results are distinct from the CWT version of db7. One obvious reason being that the DWT db7 provides an orthonormal basis of wavelets so that there is no redundancy in the representation of the transform, unlike the CWT db7.

Modeling and Metric with SVD

By themselves, the time-frequency analyses alone are not sufficient for detecting anomalies or faults. All of the time-frequency methods used in this paper, except the discrete wavelet transform, expand the dimensionality of the representation of the data. Although various researchers [1,3,4,5,7,6,20,22] have found early indications of faults with time-frequency methods, manually inspecting the results of the transforms will not produce a viable method of detecting faults in a HUMS system. Fault detection can be simplified by projecting the transformed data onto a lower dimensional model subspace. A model will be made based upon a subset of the time-frequency analysis data called the training set. This technique produces a small number of coefficients that describe the data in the model subspace and a residual that is the difference between the data and the data projected onto the model subspace. The model coefficients and/or the residual can then be used for fault detection. In this work, we will use the ratio of the rms of the residual to the rms of the transform of the data as a metric for fault detection. The value of this metric ranges from 0 for data that is perfectly described by the model to 1 for data that is orthogonal to the model, meaning that it is totally outside the model. This normalized residual will be tracked and when the level increases there is indication of a change in the state of the gear outside the range of the training set.

The basis defining the model are derived from data in the training set. By deriving the basis vectors for the model from the data rather than a standard expansion such as the Fourier series, data will fit better with fewer terms. This kind of modeling has been used to compress data for a variety of problems [33,34,35,36,37,38,39]. Several methods, Principal Component Analysis, Blind Signal Separation, Karhunen-Transform and Singular Value Decomposition accomplish this modeling with similar mathematics. The modeling is described here within the framework of the Singular Value Decomposition.

For modeling, the time-frequency analysis of each data record is reshaped into a one-dimensional column vector and all of the vectors from the transformed data set are placed into a matrix $\mathbf{X}$. The training set, $\mathbf{Y}$, is a subset of $\mathbf{X}$. $\mathbf{Y}$ is an $I \times J$ matrix of J data record transforms, I points long. $\mathbf{Y}$ is factored with a Singular Value Decomposition (SVD). The SVD produces a diagonal matrix of singular values, $\mathbf{S}$, and two unitary matrices, $\mathbf{U}$ and $\mathbf{V}$. These matrices are related by $\mathbf{Y} = \mathbf{U} * \mathbf{S} * \mathbf{V}'$. There are $K = \min(I, J)$ singular values. The columns of $\mathbf{U}$ form a basis for the column and left null space of $\mathbf{Y}$. The columns of $\mathbf{V}$ form a basis for the row and right null space of $\mathbf{Y}$. The singular values indicate the amount of each basis vector in the matrix $\mathbf{Y}$. A subset $\mathbf{W}$ is formed from the columns of $\mathbf{U}$ depending on how many of the largest singular values in $\mathbf{S}$ are considered significant. The model consists of the projection of the data onto the subset basis, $\mathbf{W}$. The coefficients, $\mathbf{D}$, from this projection now form a compact description of the data analyzed with the time-frequency method,

$$\mathbf{D} = \mathbf{W}' * \mathbf{X} \tag{10}$$

where $\mathbf{W}'$ is the Hermitian transpose of $\mathbf{W}$. Each column in $\mathbf{D}$ contains the coefficients for the model of the data in the corresponding column of $\mathbf{X}$.

A model representation of the data, $\mathbf{M}$, is then created

$$\mathbf{M} = \mathbf{W} * \mathbf{D} \tag{11}$$

Now that there is a model, define the residual, $\mathbf{R}$, to be

$$\mathbf{R} = \mathbf{X} - \mathbf{M} \tag{12}$$

Each column of $\mathbf{R} = (\mathbf{R}_1, \mathbf{R}_2, ..., \mathbf{R}_m)$ contains the residual for the corresponding column of the transformed data, $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_m)$. Define a real-valued metric as the ratio of the rms of the residual to the rms of the transformed data, where l stands for a column of data, $1 \le l \le m$

$$r(l) = \frac{\|\mathbf{R}_l\|}{\|\mathbf{X}_l\|} = \frac{\|(\mathbf{I} - \mathbf{W}\mathbf{W}')\mathbf{X}_l\|}{\|\mathbf{X}_l\|} \tag{13}$$

This metric measures how each data record differs from the model. If a data sample, $\mathbf{X}_l$, is equal to a linear combination of the columns of $\mathbf{W}$, then the data sample is in the model subset, $\mathbf{W}\mathbf{W}'\mathbf{X}_l = \mathbf{X}_l$ and $r(l) = 0$. If a data sample, $\mathbf{X}_l$, is orthogonal to $\mathbf{W}$, then the data sample is completely outside the model subset, $\mathbf{W}\mathbf{W}'\mathbf{X}_l = \mathbf{0}$ and $r(l) = 1$. If a data sample, $\mathbf{X}_l$, is equal to the sum of a linear

combination of the columns of $\mathbf{W}$ and a component orthogonal to $\mathbf{W}$, then the data sample is outside the model subset with a component in the model subset, $\mathbf{0} < \mathbf{WW}'\mathbf{X}_l < \mathbf{X}_l$ and $0 < r(l) < 1$. The more of $\mathbf{X}_l$ in the model subspace, the closer the metric is to 0. The more a data sample is like the samples in the training set, the closer the metric is to 0. When this metric is applied to the time-frequency transforms of the gear vibration data, the extracted model is built up of the local time-frequency features in the signals comprising the data set. High dimensional data have been compressed down to a single number for each data record.

For fault detection, the training set would contain a sample of data records from the gear running without faults. Measurements with no alterations to good gears were not available for this data set. In the current work, the training set contains data records from the earlier part of the test where the gear was expected to contain the least amount of damage. Models were first made using torque levels of 80% and 100% from the first 126 data records for the training set. The results showed anomalies in some of the models for the points 27-32 and 44-53. A new training set was selected to contain half the data records from 1-26, 33-43 and 54-126. The resulting model coefficients and residuals for the training set were similar to those for the non-anomalous data in the first 126 data records and not in the training set; thus, giving confidence to the model.

The number of basis vectors or modes to use in the model must be determined for each model. If too few modes are used, the model will poorly represent the transformed vibration signal from the good gear. If too many modes are used, more computation will be required producing more coefficients than needed and the model may over fit the data. Several criteria are available. The number of modes can be chosen to account for a fixed amount of rms or variance in the training set [40]. This method requires selecting a cutoff criterion. The number of modes can be chosen by examining the basis vectors and choosing only those that look like a signal [39]. This method is labor intensive and requires selecting a cutoff criterion. Another way to select the number of modes is by statistical hypothesis testing of the multiplicity of a noise eigenvalue in the singular values to distinguish between noise and signal [41,42,43]. The authors chose to use the last method that is commonly used in array signal processing for methods that require knowledge of the number of signals in the data. This method makes the assumption that data are the sum of an unknown number of stationary signals and ergodic Gaussian random noise.

When the individual data records, $\mathbf{Y}_l$, in the matrix $\mathbf{Y}$ are random samples from a process containing j signals and random noise of amplitude σ, the first j of the J singular values are estimates of the signal strengths. The rms of the remaining singular values is an estimate of the noise strength, σ. For Gaussian random noise, the remaining lower level singular values come from a normal distribution with a mean of σ. A Minimum Description Length (MDL) is found by taking a maximum likelihood estimate of the parameters. Given a vector, s, of the K singular values, the estimate of the number of signals is the value of j that minimizes the MDL

$$MDL(j) = -L(j) + \frac{j(2D - j - 1)}{2}\ln(j) \tag{14}$$

where

$$L(j) = J(K-j)\ln\left[\frac{\left(\prod_{j+1}^{k} s^2\right)^{\frac{1}{(K-j)}}}{\frac{1}{(K-j)}\sum_{j+1}^{K} s^2}\right] \tag{15}$$

In determining the number of degrees of freedom in the MDL calculation, the number of points in a data record, $D = 256$, was used, instead of $D = M$ for the data length. This was done because most time-frequency methods produce redundant information from the original shorter data record.

RESULTS

Fourteen data records were chosen as representative of what occurred during the major points of the test. Appendix A contains a table and plots of the data. Although nine of the teeth had been notched in the beginning, it was assumed that for a period of time the system would run in a less damaged state. Measurements from data records 209 and below appear to contain the least, if any, damage. The gear had severe damage at the end of the test. Measurements from data records 240 and above clearly indicate substantial damage to the gear.

The rms level is a simple magnitude of the signal in the time domain and frequency domain. Figure A1 displays the measured torque in the system and the rms level of accelerometer number 1, with the fourteen selected records marked on the plots with stars. Moderate increases occur in the rms level at records 240, 390 and 420. A large increase occurs in rms around record 440 and maximum occurs at 443. The rms amplitudes of the time-frequency transforms of the data (fig. A2 and A3) follow a different trend. The amplitudes of most of the time-frequency transforms decrease at each torque increase with the following exceptions: negligible decrease for the WV-CW transform at the 125% to 150% (records 209, 210) torque increase; small amplitude increase for the CWT at the 80% to 100% (records 43, 44) torque increase; and negligible decrease for the CWT at the 125% to 150% torque increase. Table A1 lists all of these magnitudes for the fourteen selected data records.

Data and Transforms

Time History

Figure A4 shows the synchronously averaged time history for the 14 selected data records. The 19 per rotation gear mesh signal is clearly present in the earlier time histories for most of the data records in the two lowest torque levels (80% and 100%) up to record number 126. In a few cases, such as record 44, some of the gear mesh waveforms might be merging to present fewer than 19 distinct waveforms. For time histories from records with torque level of 125% (records 127 through 209), there is less clear definition of the 19-per rotation periodic signal between 0 and 0.2 of the rotation and between 0.7 and 1. Signals from adjacent teeth may be starting to merge. The amplitudes of the signals are very similar up through about record 236. The signal changes dramatically in amplitude and shape between records 236 and 240. The amplitude of the signal is much larger in records 240 and above, and sometimes shows considerable variation with rotation

angle. The 19 per rotation gear mesh waveform is no longer recognizable in record 240 and above. Records 240 through 285 contain strong waves near the start and end of the rotation with very weak waves in the middle. The shape changes to contain 9 distinguishable waves from records 394 through the end of the test. Time histories from the last 9 data records (441 through 449) have the largest amplitudes and change shape notably from record to record. The shape of record 443 resembles a single frequency sinusoidal wave.

Amplitude Spectra

Figure A5 shows the amplitude of the Fourier Transform of the synchronously averaged time history of the 14 selected records. The signal displays the 19 per rotation gear mesh component as the dominant frequency in the early data up through record 209. This gear mesh component is significant up through record 236 and remains visible through all the data, although becoming insignificant beyond record 394. In comparison, the gear mesh frequency is not recognizable in the Time History for data records 240 and above. Three or four harmonics of the gear mesh component are visible throughout the three lower torque levels (80%, 100% and 125%), up through record 209. While running the gear with the highest torque level (150%), the gear mesh harmonics are clearly visible in the earlier records, then abruptly diminish between records 236 to 240, and are insignificant after that. As the test proceeds, energy moves to frequencies lower than the gear mesh harmonic. By the end of the third torque level (125%, record 209), the spectra contain significant energy around 9 and 10 per rotation of the gear. By record 240, most of the energy is concentrated in a broad range from 7 to 12 per rotation of the pinion. In the final data records, 440 to 449 (150% torque) the energy in concentrated in the 7 and 9 per rotation frequencies.

Short Time Fourier Transform

Figure A6 shows the amplitude of the STFT for the 14 selected data records. The time shift in the STFT varies along the horizontal axis over one rotation of the pinion. The vertical axis spans the frequency range with the lower frequencies at the top of the axis and the higher frequencies near the bottom. The block size represents the window size that was chosen. In this case, there are 29 blocks that span the entire STFT. Unlike the CWT, this is the smallest resolution because of the fixed window size. The smallest frequency is 1/8 of the rotation or 2.375 times the gear mesh frequency; the gear mesh frequency will register at indices 3 and 4. With this resolution, the data span 17 frequency indices instead of 128.

In the early records with torque of 80% and 100%, the gear mesh signal shows as a dark band at indices 3 and 4 at the gear mesh frequency. Variation in the level with rotation can be seen. Record 44 shows greater variation with the amplitude reduced near 0.8 rotation. Examination of the corresponding time history in figure A4 also displays lower signal amplitude near 0.9 rotation of the pinion. In the 125% torque region (records 127-209), the amplitude for the gear mesh harmonic becomes more variable and the gear mesh frequency component no longer always contains the largest amplitude; at some positions in the rotation, a lower frequency has the largest amplitude. By record number 240, the largest amplitude always occurs at a lower frequency. The rotation position with the largest amplitude moves around through the remainder of the test. The amplitude of single components increase around record 240 with a maximum level at record 443.

Choi-Williams Distribution

Figure A7 shows the Wigner-Ville distribution with the Choi-Williams kernel. The time shift in the Choi-Williams Distribution varies along the horizontal axis over one rotation of the pinion. The resolution in the time shift is the same as the resolution in the data, 256 points per rotation. The vertical axis spans the frequency range with the lower frequencies at the top of the axis and the higher frequencies near the bottom. Since the Choi-Williams data is four times finer than the Fourier transform, the index for the gear mesh frequency is at four times 19, or at 76.

In the data records for the 3 lower torque levels (records 1 through 209), the signal of the gear mesh frequency component shows as a dark line near index 76 throughout the rotation and as less prominent vertical lines spaced at 19 per rotation. These features are present, but no longer dominate in records 210 through 236 (150% torque). These features are insignificant or not even visible in records 240 and above. Dark horizontal lines from the second and third harmonic of the gear mesh frequency are present in varying degrees up through record 236. Energy shows as dark horizontal lines at frequencies below the gear mesh frequency starting with record 127, through the end of the test. The lower frequency energy is apparent between rotation 0.6 to 0.8 in records 127 through 209 and is stronger in records 210 through 236. A vertical line is prominent in that region on record 210, indicating a wide frequency event in a short span of time. The corresponding time history in figure A4 shows a sharp event near 0.7 rotation. Low frequency energy is strong near rotation 0.1 in records 236, 240 and 285 and diminishing in relative importance for later records. Low frequency energy is strong near rotation 0.9 in records 240, 285, 394 and 415. A closer look at records 240 and 285 shows multiple lines at low frequency with relative maxima at different frequencies for different rotation positions of the gear. A single frequency component is very strong in record 443. The corresponding time history and amplitude spectra (figures A4 and A5) also display a strong frequency component. Two low frequency components are strong in the last record, 449. In general, the amplitude of individual components is low in the early part of the test, up through record 236. The amplitude increases gradually up to about record 440. Above record 440, the amplitude is large, with a local maximum at record 443.

Continuous Wavelet Transform

Figure A8 shows the Continuous Wavelet transform for the 14 chosen records. The time shift in the CWT varies along the horizontal axis over one rotation of the pinion. The resolution in the time shift is the same as the resolution in the data, 256 points per rotation. The vertical axis represents each scale of the wavelet. Low scale and high frequency appear at the top while high scale and low frequency appear at the bottom of each plot. Scale number 11 maximizes the gear mesh frequency. The value at each point of the image represents the transformed data at that scale. Notice, for example, how in record 27 the amplitude oscillates at 19 per rotation in the region around the scale that maximizes the gear mesh frequency. Also, notice how at the edges of the image the amplitude is less than it is at the middle of the image. This corresponds to the time data in which the amplitude of the data also increases in the middle of the rotation.

The gear mesh signal is the dominant feature in the transforms for the two lower torque levels (80% and 100%, 1 through 126). A lower frequency wave (higher scale) near 0.7 rotation is visible in records 127 through 210 along with the gear mesh signal. A low frequency wave is apparent near 0.1

rotation in records 236 through 285. A relatively steady low frequency wave is apparent in records 394 through the end. The amplitude of wavelet components is relatively low up through record 236 and largest for record 443.

Discrete Wavelet Transform

Figure A9 shows the Discrete Wavelet transform for the 14 chosen records. The DWT is distinct from the other transforms in this paper. It does not expand the dimensionality of the representation of the data. In this case, where seven orders have been used, there are seven vectors generated, which are referred to as details, and one vector that is referred to as an average. In figure A9, the details with the highest frequencies are at the right side of the plot. The frequency range covered by a detail increases from the lowest in detail 7 near the left side to the highest in detail 1 on the right side of the plot. The average vector, which contains the lowest frequency, is at the left side of the plot. In figure A9, the time shift in the DWT covers one rotation of the pinion for each detail and the average. The time resolution is finer in the details at lower scale and higher frequency near the right side of the plots. The gear mesh frequency at 19 shaft orders does not match any of the scales. Signals at the gear mesh frequency are split between details at scales 3 and 4.

In the early data records (1 through 209, 80%, 100% and 125% torque), most of the energy is in the detail at scales 3 and 4, as expected for the gear mesh frequency. More energy shifts to detail at scale 5, a lower frequency, by record 236. For records 240 through the end, most of the energy is in the detail for scales 4 and 5. As with the other transforms, the amplitude is highest for record 443.

SVD Models of Time-Frequency Analyses

SVD models were made for the time history, the real valued amplitude spectra, the complex valued Discrete Fourier Transform of the Time History, and all four time-frequency transforms. The training set for all seven models consisted of half of the data in records 1-26, 33-43 and 54-126. All training set data came from the two lowest torque levels.

Appendix B contains more detailed information on the residuals for the 14 data records. Table B1 lists the normalized rms residual. Figure B1 shows the normalized rms residual for the SVD models of the Time History, Amplitude spectrum, Fourier Transform of the Time History and the DWT. First, notice that only two curves are visible on the plot. The normalized rms residual for the models made from the Time History, Fourier Transform of the Time History and the DWT all coincide to within about 10^{-13}; these three models appear to be the same model. The Time History, Fourier Transform and DWT can all be transformed into each other. That is, each contains exactly the same information about the data. The eigenvectors found in the SVD models of these three descriptions of the data also match each other in the sense than any one can be transformed into another with an error of only 10^{-13} in the computational environment. The normalized rms residual for the Amplitude Spectra differs from and lies below the residual for the other three descriptions. The Amplitude Spectra contains no phase information, which means it contains less information than the Time History. The levels of the residuals are approximately constant and low for each of these four models of the 80% and 100% torque levels except for a small increase for the records 44-53 on the models made from the Time History, Fourier Transform and DWT. The levels jump up to a higher

level at the change to 125% torque (record 127) and again at the change to 150% torque (record 210). For the fourth torque level of 150% in data records 210-449 the residuals for all transforms initially jump up at data record 210 then rise steadily to record 238, then jump up again and remain at a high level for the remainder of the data records. In this last region, small changes occur in the residual occurring around records 300, 410 and 440. All residuals above data record 240 are at high levels, close to one. These high levels indicate that the shape of the vibration signals differ greatly from the shape of those in the training set. For example, the time histories in figure A4 records 240,285, 394, 415, 443 and 449 do not resemble those in records 27, 28, 44 and 126.

The normalized rms residuals for the SVD models of the four Time-Frequency transforms (fig. B2) are very similar to the residual of the Time History model (fig. B1). Differences in qualitative behavior occur in three regions: records 27-32 (80% torque); records 44-53 (100% torque); and records 127 – 209 (125% torque). First, for the 80% torque level, the residual level is at the low baseline level for all time frequency methods except the CWT. For data records 27-32 the residual increased almost 100% over baseline for the CWT. Second, for the 100% torque level, the residual levels are in two ranges, the low baseline and a higher level. For data records 44-53, the residuals for the STFT, WV-CW and DWT show a slight increase above the baseline, the residual for the CWT show 100% increase. Third, for the 125% torque level, the residual for the STFT gradually increases, the residual for WV-CW initially jumps up to about 0.5 then decreases to about 0.3 for most of the records, the residual for the CWT initially jumps up then decreases then steadily increases and the residual for the DWT remains about constant.

Figures B3 through B8 show the residuals for the Time History, Amplitude Spectrum, Fourier Transform, STFT, WV-CW, CWT and DWT respectively. In all of these figures, the residual resembles noise in the data records where the data fits the model well. Departure from noise shows where the data do not fit the empirical model. The residuals look very much like the data in records 240 and above.

Residuals show how the data differ from the empirical model derived from the training set. Examination of residuals in the anomalous records, 28 through 32 and 44 through 53, and in the transitional records 127 through 230 will reveal how these records differ from the empirical model. For record 28, the residual Time History (fig. B3) and CWT (fig. B7) both show a low frequency (3 per rotation) wave, and the DWT (fig. B8) contains a high level in detail 6. For record 44, the residual of the CWT (fig. B7) shows a less distinct low frequency wave and the DWT (fig. B8) shows more energy in detail 3. For record 127, the residual of the Amplitude Spectrum (fig. B4) shows non-random harmonic content. For record 209, the residual of the Amplitude Spectrum (fig. B4) shows energy below the gear mesh harmonic and the CWT (fig. B7) shows stronger waves near rotation 0.8. For the residual of record 210, the Time History (fig. B3) contains spikes near rotation 0.7, the Amplitude Spectra (fig. B4) contains low frequency energy and deviation near the gear mesh frequency, and the STFT, WV_CW and CWT all show waves near rotation 0.7. For the residual of record 236, the Amplitude Spectra contains much low frequency energy and the STFT, WV-CW and CWT all show low frequency waves near rotation 0.1 and 0.9. The residuals for records 240 and above appear very much like the data.

The coefficients generated by the model and not shown here relate to both the torque level of the gear and the health of the gear. In all of the models, the coefficients change when the torque

changes. The coefficients also distinctly change for the seven models at data records 27, 33, 44, 54, 240, 280, 393, 408, 440 and 443

DISCUSSION

Computer Resources

All of the computations were done on a generic 550 MHz Pentium III Xeon with 1 GB of memory running Redhat Linux 6.2.

Referring to Table 2, in increasing order, a single STFT took .005 seconds to complete, a DWT took .096 seconds and a CWT took 1.37 seconds. The WV-CW, which took 2.48 seconds to perform, was by far the longest single method.

Table 2. Method Comparisons

	Time (sec.)	SVD (sec.)	Samples per rev.	Spectrum shaft harmonics
Short Time Fourier Transform	0.005	6.0	256	128
Choi-Williams	2.48	176.4	256	64
Continuous Wavelet	1.37	77.9	256	128
Discrete Wavelet	0.096	4.7	256	128

In addition, in increasing order, the SVD times for a single DWT took 4.7 seconds. A STFT took 6.0 seconds and a CWT took 77.9 seconds. Again, the method that took the longest was the WV-CW, taking 176.4 seconds.

Although there were no exact figures kept on the amount of space used for each of the methods, from estimates, the DWT took the least amount of space. The STFT took the second most, and the CWT took the next. The WV-CW, with a very large margin, required the most memory.

The WV-CW method took the most computer resources, both time and space, without showing a clear benefit over other methods.

SVD Modeling and Metric

The Time History, Amplitude Spectrum, Fourier Transform and four Time-Frequency analyses clearly show changes in the data over the course of the experiment. The experiment did not provide

a way to directly determine the state of the gear during the experiment. The gear contained notches on nine of the teeth before data record one and the gear sustained considerable damage by the end of the experiment. The damage probably occurred in stages. The modeling of the time-frequency transforms at data records 27, 33, 44, 54, 127, 210, 240, 285, 400, 410, 440 and 443-449 indicates changes in the vibration signal. Many of the changes are also observed in the simpler models based upon the time history, amplitude spectra and FFT.

The local increases in the rms residual in data records 27-32 and 44-53 are very evident on the models for the CWT models, but not clear in the rms residual for the other the time-frequency transforms nor in the SVD models made from the time and frequency domains. The model coefficients in these local regions differ for the models of all of the time-frequency analyses. There was no reason found for changes in the rms residual of the CWT model compared to the other TF methods in these two local regions. Perhaps the CWT is detecting something in the gears or the rig that the other methods are not, but this could not be corroborated.

The large jump in the residual at record 127 is probably a combination of change to a torque level not included in the training set and change in the shape of the time history due to minor damage to some of the gear teeth. The residual change indicates a change in the character of the signal, not the amplitude of the signal. The time history in figure A4 for data record 127 shows some of the waveforms with a tooth-width wavelength beginning to merge near 0.1, 0.75 and 0.85 of the rotation. In comparison the time histories from the earlier records more clearly contain 19 individual waveforms for the 19 teeth on the pinion. With minor damage to a gear tooth, the tooth may no longer carry its share of the load thus producing changes in the vibration signal, including the merging of waveforms. It is plausible that if a gear tooth contains a small defect, the vibration signal due to that tooth will differ from the vibration signal due to good teeth and the difference will increase with increasing load on the tooth. The large jump at the second torque change may occur because the torque is increasing from 100% to 125% of full load. A training set containing data with a good gear running at 125% torque is needed to make more definitive statements. The gradual increase in the residual for the STFT and CWT models in the third torque region may indicate gradually increasing damage (fig. B2).

The steep jump in the residual at the final torque increase (data record 210) and gradual increase for about the next 40 data records can be interpreted in the same manner as the jump and increase after the second torque change. The time history in figure A4 for data record 210 shows even more merging of waveforms on the scale of the gear-tooth to a longer wavelength near rotation 0.1, 0.75 and 0.9.

All changes after the final torque definitively indicate changes in the state of the gear. A steep jump in residual and abrupt changes in the coefficients of the models occurred near data record 240 indicating an abrupt change in the gear. Gear teeth probably fractured near data record 240. Changes indicating discrete changes in the gear condition also occurred near data records 285, 400 and 410.

For the last 9 data records, the coefficients of the models show more drastic change than the residual. Many of the coefficients increase by a factor of 2 or more over a few data records then decrease to low levels while the residual increases or remains high. Very severe damage probably

occurs over the last several data records. The authors believe the input shaft failure caused the large changes observed in the last few data records.

Residuals from models of the different time-frequency analyses have some different characteristics. The CWT indicated sensitivities that the other TF methods did not. Although not currently identified, they might prove to be useful. The WV-CW model gave the closest fit to the data in the training set and its residual rose more than for other models. When the level of the residual from modeling indicates damage may exist in machinery, more information can be obtained in the modeling by further examination of the residual. If the damage is of a local nature, the time-frequency representation of the residual will show structures localizing the damage. Such structures were observed in the measurements of current test waveforms changing size from 1/19 of a rotation to about 2/19 of a rotation. Examination of such residuals from different kinds of faults in different machinery might lead to the ability to classify the damage.

The time-frequency analyses distinguish events at distinct times, so signals used in the modeling must be consistently synchronized. When the interrupter signal used for synchronization is on the same shaft as the gear being analyzed, there is no alignment problem. When the phase alignment queue is from a shaft spinning at a different frequency, something must be done to account for the proper phase. One option is to align the data with a circular correlation correction. Another option is to use synchronously averaged data from all possible phase alignments of the gear in the training set. Of the seven data forms reported here, only the Amplitude Spectra could be used without accurate phase alignment.

The vibration signature of a gear depends upon the torque load on the gear. The amplitude and the relative harmonic content or signal shape are both known to change with changing load. To properly model the vibration, these signal amplitude and shape effects must be included in the model. In the current study, two discrete torque levels, 80% and 100%, were included in the one empirical model. An alternative empirical model could be made by creating a distinct model for each torque level, utilizing measurements from the gear in good condition to train a model for each torque level. Individual models for each torque level may work well for the test rig environment; however, in flight, the torque range is broad and continuous. A modeling method that incorporates the nonlinear dependence of the vibration signal on torque, and possibly other operating parameters, is needed. A nonlinear extension of Principal Component Analysis [44] (an alias for SVD) might accomplish this.

The SVD modeling of the transformed data shows promise for development into an automatic fault detector for use on machinery. Extensive training and test sets covering all operating conditions of the gear are needed. The distribution of the normalized rms residual can then be determined by the distribution found for the training and test sets. With the distribution the trigger level on the normalized rms residual can be set to yield a given false alarm rate. The sensitivity of the fault detection can be investigated with test rig data of gears with and without faults. The false alarm rate can be investigated with measurements made in flight. This fault detection will have the combined benefits of sensitivity to faults that produce changes in the time-frequency analyses and the ability to automate fault detection with a single number metric.

CONCLUSION

The Short Time Fourier Transform, the Choi-Williams modified version of the Wigner-Ville distribution, the Continuous Wavelet Transform and the Discrete Wavelet Transform all reveal changes in the vibration measurements of a damaged spiral bevel pinion gear.

The Choi-Williams modified version of the Wigner-Ville distribution and the Continuous Wavelet Transform both use considerably more computational resources than the Short Time Fourier Transform and the Discrete Wavelet Transform.

A single number metric applicable to automatic fault detection was introduced that can be produced from any systematic numerical representation of the vibration signals. Vibration measurements of the gear operating without faults are needed to generate the model using the metric.

The new metric revealed sensitivities with the Continuous Wavelet not apparent in the other TF methods, thus indicating the additional computational resources needed for the Continuous Wavelet Transform perhaps add value in earlier detection.

More testing of the new metric is needed to determine its value for automatic fault detection and to develop methods of setting the threshold for the metric.

REFERENCES

[1] Staszewski, W. J., Worden, K., Tomlinson, G. R., "Time-Frequency Analysis in Gearbox Fault Detection Using the Wigner-Ville Distribution and Pattern Recognition," Mechanical Systems and Signal Processing, Vol. 11, (5), 1997.

[2] Zakrajsek, James J., Handschuh, Robert F., Lewicki, David G., Decker, Harry J., "Detecting Gear Tooth Fracture in a High Contact Ratio Face Gear Mesh," NASA Technical Memorandum 106822, 1995.

[3] Hambaba, A., Huff, E. and Kaul, U., "Detection and Diagnosis of Changes in the Time-Scale Eigenstructure for Vibrating Systems," Aerospace Conference, 2001, IEEE Proceedings. , Volume: 7, 2001.

[4] Samuel, Paul D., Pines, Darryll J., Lewicki, David G., "A Comparison of Stationary and Non-Stationary Metrics for Detecting Faults in Helicopter Gearboxes," Journal of the American Helicopter Society, Vol. 45 (2), April 2000.

[5] Wang, W. J., McFadden, P. D., "Early Detection of Gear Failure by Vibration Analysis--I. Calculation of the Time-Frequency Distribution," Mechanical Systems and Signal Processing, Vol. 7, (3), 1993.

[6] Williams, W. J., Zalubas, E. J., "Helicopter Transmission Fault Detection Via Time-Frequency, Scale and Spectral Methods," Mechanical Systems and Signal Processing, Vol. 14, (4), 2000.

[7] Wang, W. J., McFadden, P. D., "Early Detection of Gear Failure by Vibration Analysis--II. Interpretation of the Time-Frequency Distribution Using Image Processing Techniques," Mechanical Systems and Signal Processing, Vol 7, (3), 1993.

[8] Loughlin, P., and Cakrak, F., "Conditional Moments Analysis of Transients with Applications to Helicopter Fault Data," Mechanical Systems and Signal Processing, Vol 14, (4), December 1999.

[10] Oehlmann, H., Brie, D., Tomczak, M., and Richard, A., "A Method for Analyzing Gearbox Faults Using Time-Frequency Representations," Mechanical Systems and Signal Processing, Vol. 11, (4), October 1996.

[9] Choy, F. K, Huang, S, Zakrajsek, J. J., Handschuh, R., F., and Townsend, D. P., "Vibration Signature Analysis of a Faulted Gear Transmission System," NASA TM106623, June 1995.

[11] Baydar, N., and Ball, A., "Detection of Gear Deterioration Under Varying Load Conditions by Using the Instantaneous Power Spectrum," Mechanical Systems and Signal Processing, Vol. 14, (6), November1999.

[12] Boulahbal, D., Golnaraghi, M. Farid, and Ismail, F., "Amplitude and Phase Wavelet Maps for the Detection of Cracks in Geared Systems," Mechanical and Signal Processing, Vol. 13, (3), October 1998.

[13] Lin, Jing, and Qu, Liangsheng, "Feature Extraction Based on Morlet Wavelet and its Application for Mechanical Fault Diagnosis," Journal of Sound and Vibration, Vol. 234, (1), January 2000.

[14] Lopez, Jose E., Tenney, Robert R., and Deckert, James C.," Fault Detection and Identification Using Real-Time Wavelet Feature Extraction," Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, 1994.

[15] Tenney, Robert R, Deckert, James C. and Rhenals, Alonso E., "Wavelts for Helicopter Transmission Fault Detection and Classification," MFPT Proceedings No. 48, 1994.

[16] Lopez, Jose E., Deckert, James C. and Tenney, Robert R., "Improved Fault Identification using Multisensor Wavelet Based Differential Features," MFPT Proceedings No. 49, 1995.

[17] Lopez, Jose E., Yeldham, Inna A. Farber and Oliver, Kevin, "Overview of Wavelet/Neural Network Fault Diagnostics Methods Applied to Rotation Machinery," MFPT Proceedings No. 50, 1996.

[18] Ferlez, Robert, and Lang, Derek C., "Gear-Tooth Fault Detection and Tracking Using the Wavelet Transform," Proceedings of Prognosis of Residual Life of Machinery and Structures, MFPT, 1998.

[19] Ferlez, Robert J. and Lang, Derek C., "Statistical Characterization of Continuous Wavelet Transform Magnitude Feature Data for Gear-Tooth Fault Detection Thresholds," MFPT Proceedings No. 53, 1999.

[20] Dalpiaz, G., Rivola, A, and Rubini, R., "Effectiveness and Sensitivity of Vibration Processing Techniques for Local Fault Detection in Gears," Mechanical Systems and Signal Processing, Vol. 14, (3), 1999.

[21] Lin, S. T. and McFadden, P. D., "Gear Vibration Analysis by B-Spline Wavelet-Based Linear Wavelet Transform," Mechanical Systems and Signal Processing, Vol 11, (4), April 1997.

[22] Wang, W. J., McFadden, P. D., "Application of Wavelets to Gearbox Vibration Signals for Fault Detection," Journal of Sound and Vibration, Vol. 192, (5), 1996.

[23] Lin, J. and Qu, L, "Feature Extraction Based on Morlet Wavelet and its Application for Mechanical Fault Diagnosis," Journal of Sound and Vibration, Vol 234, (1), 2000.

[24] Paya, B. A., and Esat, I. I., "Artificial Neural Network Based Fault Diagnostics of Rotating Machinery Using Wavelet Transforms as a Preprocessor," Mechanical Systems and Signal Processing, Vol. 11, (5), March1997.

[25] Wang, W. J., and McFadden, P. D., "Application of Orthogonal Wavelets to Early Gear Damage Detection," Mechanical Systems and Signal Processing, Vol. 9, (5), May1995.

[26] McFadden, P. D., "Detection of Gear Faults by Decomposition of Matched Differences of Vibration Signals", Mechanical Systems and Signal Processing, Vol. 14, (5), May2000.

[27] Pryor, A. H., Mosher, M., and Lewicki, D. G., "The Application of Time-Frequency Methods to HUMS", American Helicopter Society 57th annual Forum, Washington, D.C., May 9-11, 2001.

[28] Lewicki, David G., Coy, John J., "Vibration Characteristics of OH-58A Helicopter Main Rotor Transmission," NASA Technical Paper 2705, 1987.

[29] Kaiser, Gerald, A Friendly Guide to Wavelets, Birkhauser, Boston, 1994.

[30] Choi, H. and Williams, W. J., "Improved Time-Frequency Representation of Multicomponent Signals Using Exponential Kernels," IEEE Trans. on Acoustics Speech and Signal Processing, Vol. 37, (6), June 1989.

[31] Cohen, Leon, "Time-Frequency Distributions-A Review," Proceedings of the IEEE, Vol. 77, (7), July1989.

[32] Daubechies, Ingrid, Ten Lectures on Wavelets, SIAM, Philadelphia, Pennsylvania 1992.

[33] Stapleton, John C. and Bass, Steven Craig, "Synthesis of Musical Tones Based on the Karhuen- Transform," IEEE Transaction on Acoustics, Speech and Signal Processing, Vol 36, (3), March 1988.

[34] Kirby, M. and Sirovich, L., "Application of the Karhuen- Procedure for the Characterization of Human Faces," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, (1), January 1990.

[35] Beyerbach, Daniel and Nawab, Hamid, "Principal Components Analysis of the Short-Time Fourier Transform," Proceedings International Conference on Acoustics, Speech and Signal Processing, Vol. 3, April 1991.

[36] Newab, Hamid, Beyerbach, Daniel and Dorken, Erkan, "Principal Decomposition of Time-Frequency Distributions," IEEE Transactions on Signal Processing, Vol. 4, (11), November1993.

[37] Hung, Elmer and Zhao, F., "Diagnostic Information Processing for Sensor-Rich Distributed Systems," Poc. 2nd International Conference on Information Fusion, Sunnyvale, CA, 1999.

[38] Tumer, Irem Y., Wood, Kristin L., and Busch-Vishniac, Ilene J. "A Mathematical Transform to Analyze Part Surface Quality in Manufacturing," Journal of Manufacturing Science and Engineering, Vol. 122 (1), February 2000.

[39] Penev, Penio S. and Sirovich, Lawrence, "The Global Dimensionality of Face Space," Proceedings of Forth IEEE International Conference on Automatic Face and Gesture Recognition, March 2000.

[40] Jolliffe, I. T., Principal Component Analysis, Springer-Verlag, New York, 1986.

[41] Wax, Mati and Kailath, Thomas, "Detection of Signals by Information Theoretic Criteria," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 33, (2), April 1985.

[42] Williams, Douglas, B., "Counting the Degrees of Freedom When Using AIC and MDL to Detect Signals," IEEE Transactions on Signal Processing, Vol. 42, (11), November 1994.

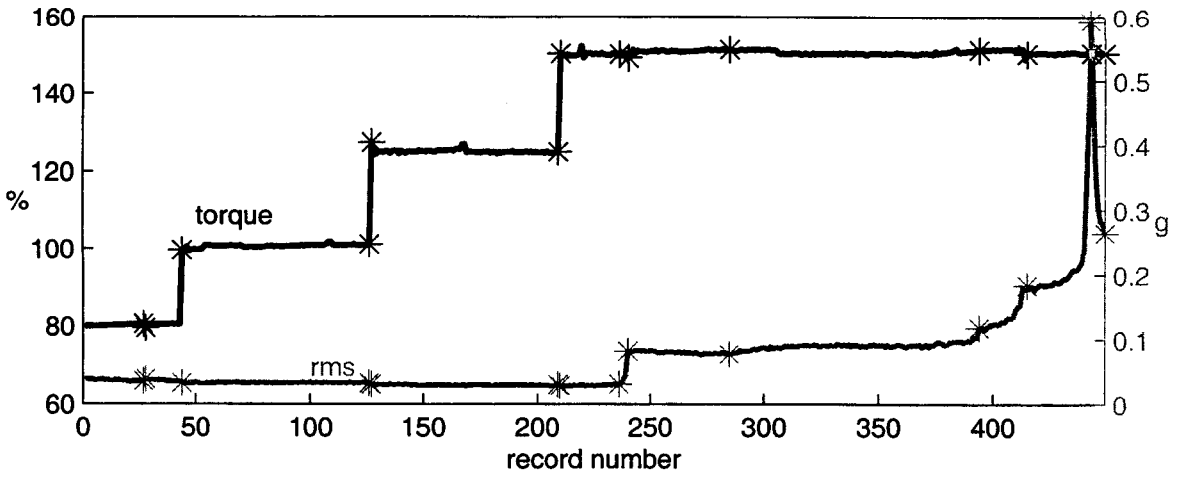
[43] Johnson, Don H. and Dudgeon, Dan E., Array Signal Processing Concepts and Techniques, Prentice Hall, New Jersey, 1993.

[44] Kambhatla, Nanda, and Leen, Todd K., "Fast Non-Linear Dimension Reduction "in Advances in Neural Information Processing Systems 6,Morgan Kaufmann Publishers, San Francisco, California 1994.

APPENDIX A. DATA

Table A1. Data from accelerometer measuring vibration of pinion gear.

record #	torque	rms	amp STFT	amp WV-CW	amp CWT	amp DWT
27	80.512	0.03494	0.52983	6.0764	4.7293	0.52983
28	79.4556	0.037836	0.57005	6.786	4.5506	0.57005
44	99.5359	0.032211	0.47025	4.7109	4.6216	0.47025
126	101.1185	0.031962	0.48194	4.9819	4.8729	0.48194
127	127.4756	0.030198	0.39148	2.3613	3.7123	0.39148
209	125.0637	0.029741	0.38353	2.8926	3.7783	0.38353
210	150.2112	0.027922	0.29745	2.0008	3.6323	0.29745
236	150.0699	0.030552	0.29666	1.9145	4.0878	0.29666
240	149.1962	0.081502	0.3137	3.1713	6.217	0.3137
285	151.2503	0.07716	0.34475	3.0619	7.1658	0.34475
394	151.1303	0.1179	0.26215	3.355	5.5059	0.26215
415	150.1734	0.18344	0.25895	7.5752	5.2841	0.25895
443	150.3467	0.59269	0.697	51.3963	21.5556	0.697
449	150.253	0.26414	0.25384	15.1243	3.2135	0.25384



FigureA1. Measured torque and rms level of accelerometer.

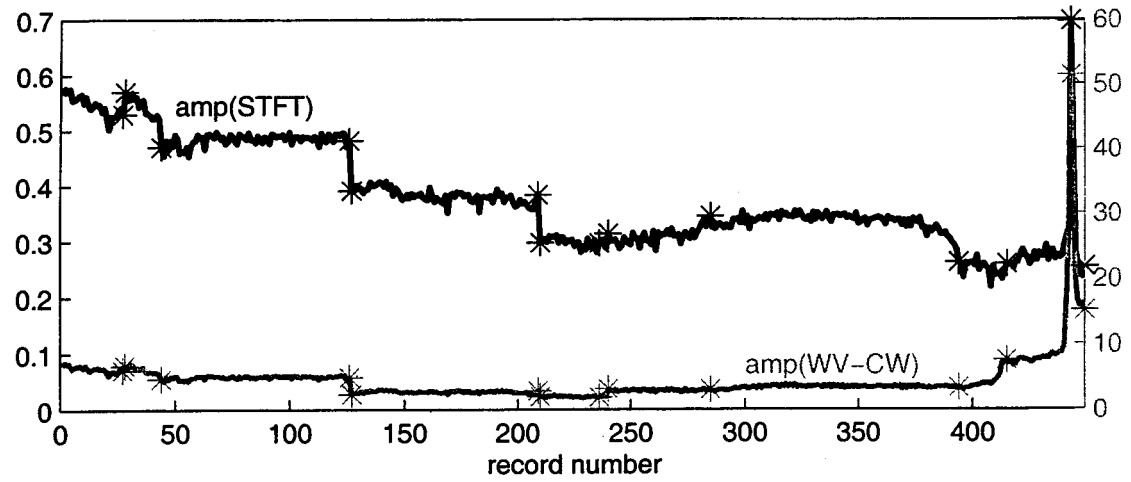


Figure A2. Amplitude of the Short Time Fourier Transform and the Wigner-Ville distribution with Choi-Williams kernel.

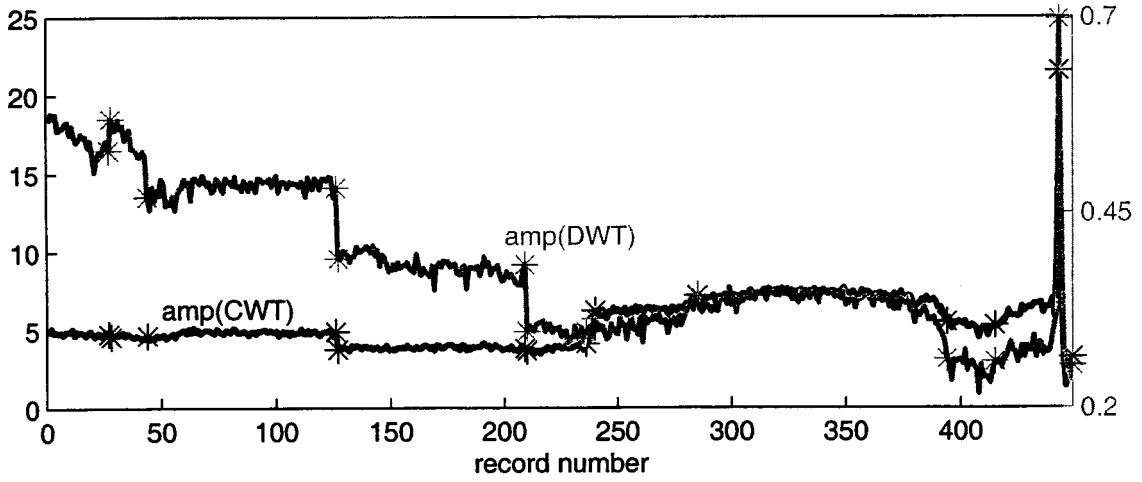


Figure A3. Amplitude of the Continuous Wavelet Transform and the Discrete Wavelet Transform.

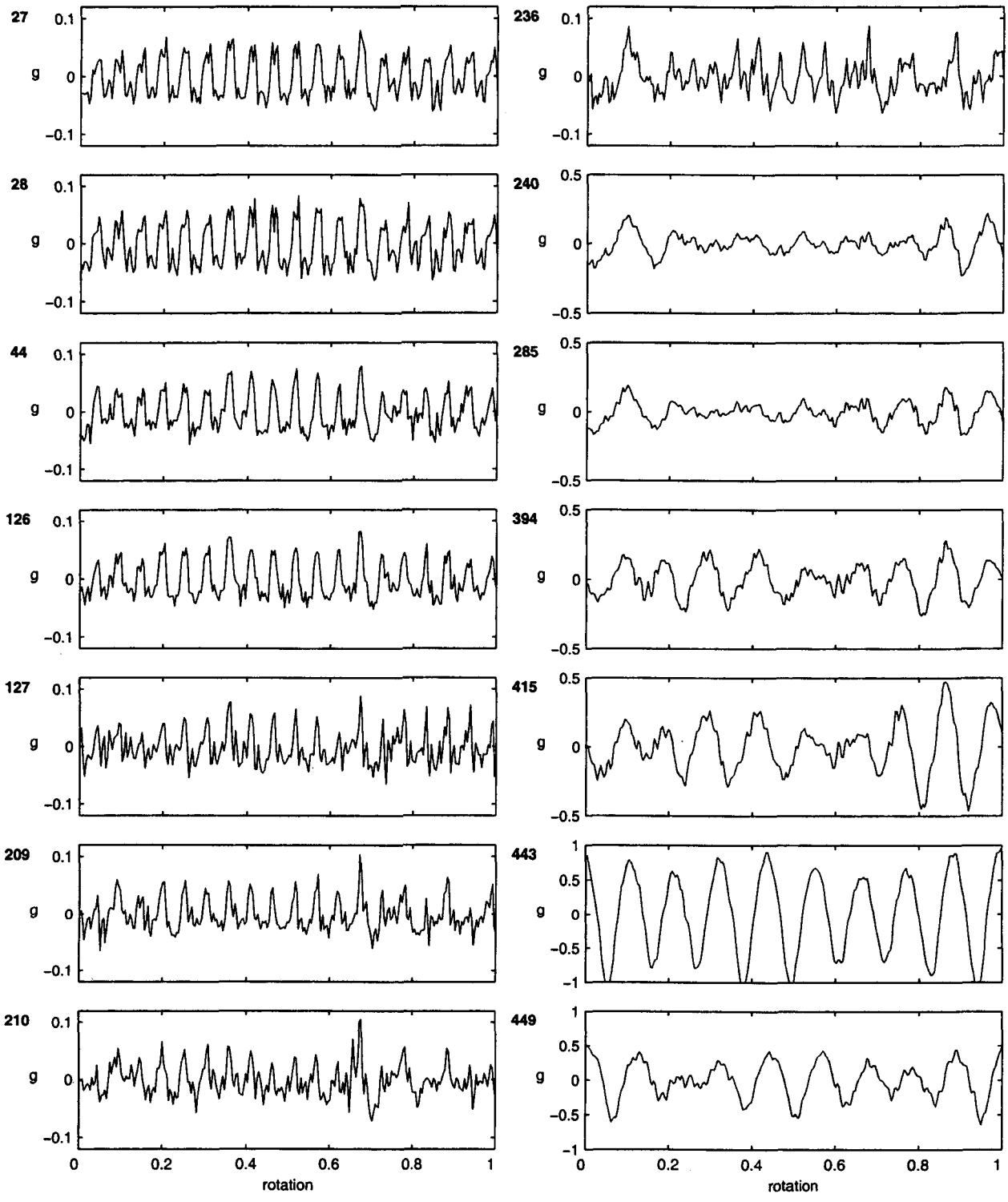


Figure A4. Time synchronous average for one rotation of gear: Data from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

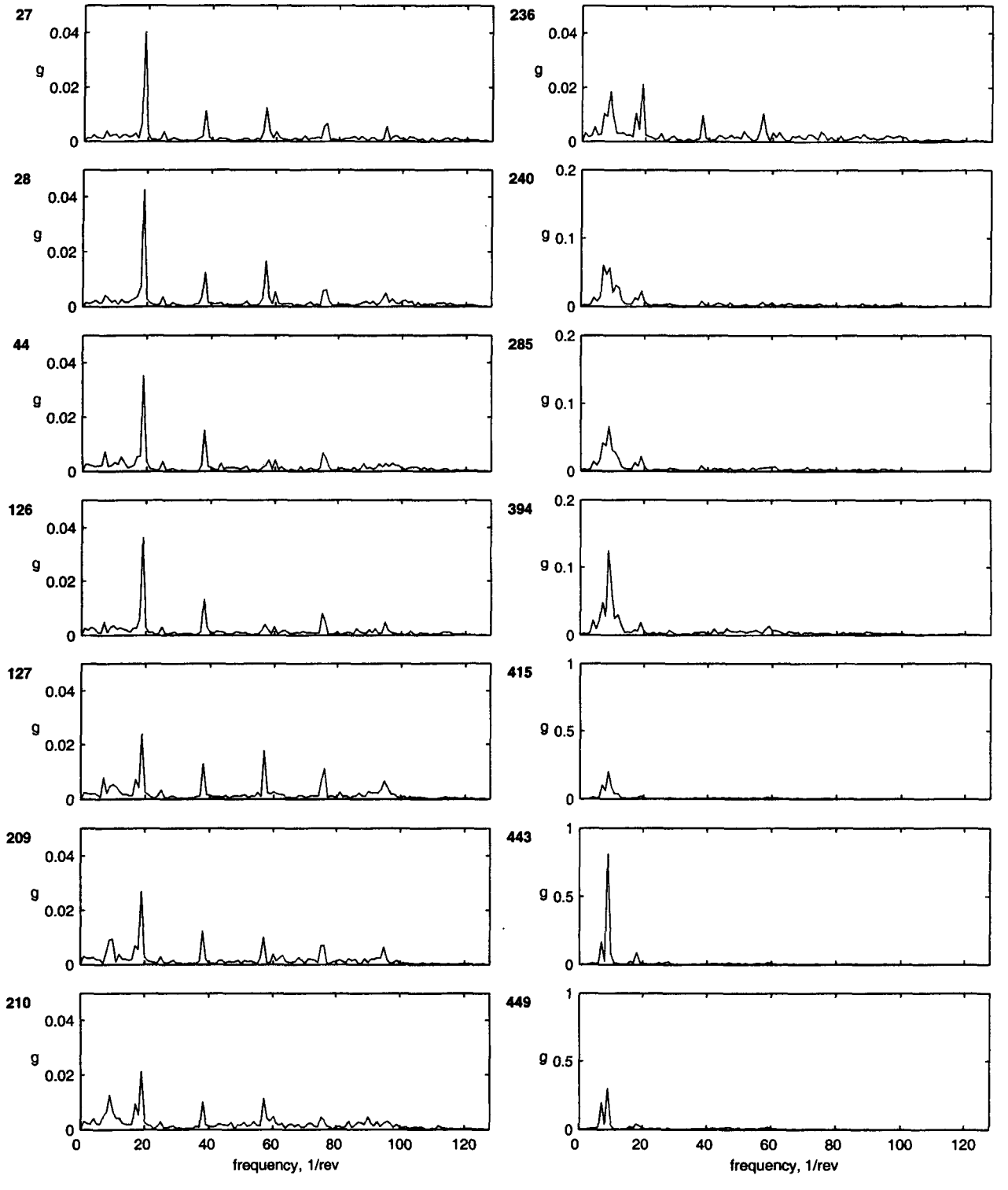


Figure A5. Amplitude of Fourier Transform of time synchronous average for one rotation of gear: Data from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

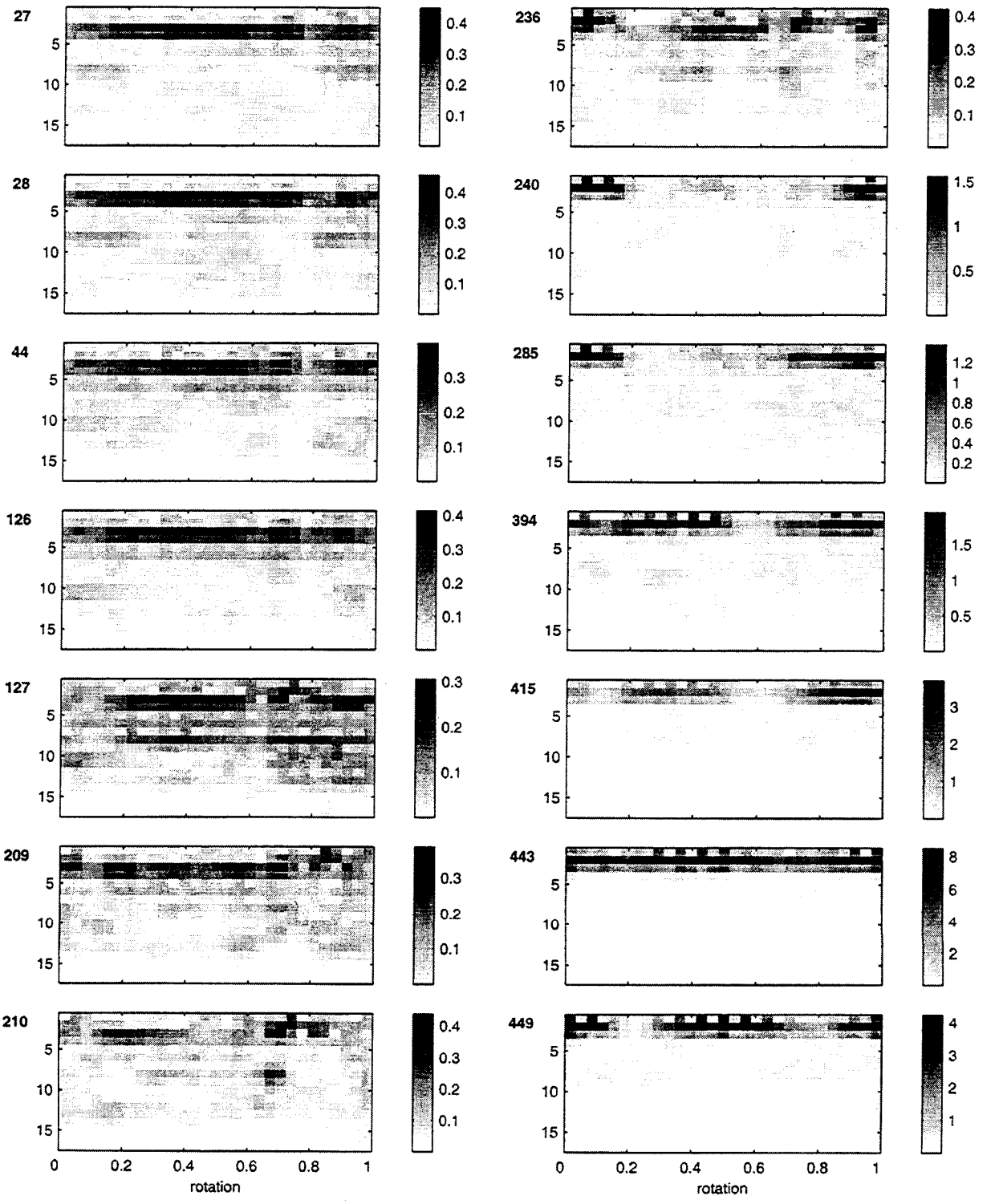


Figure A6. Short Time Fourier Transform of time synchronous average for one rotation of gear: Data from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

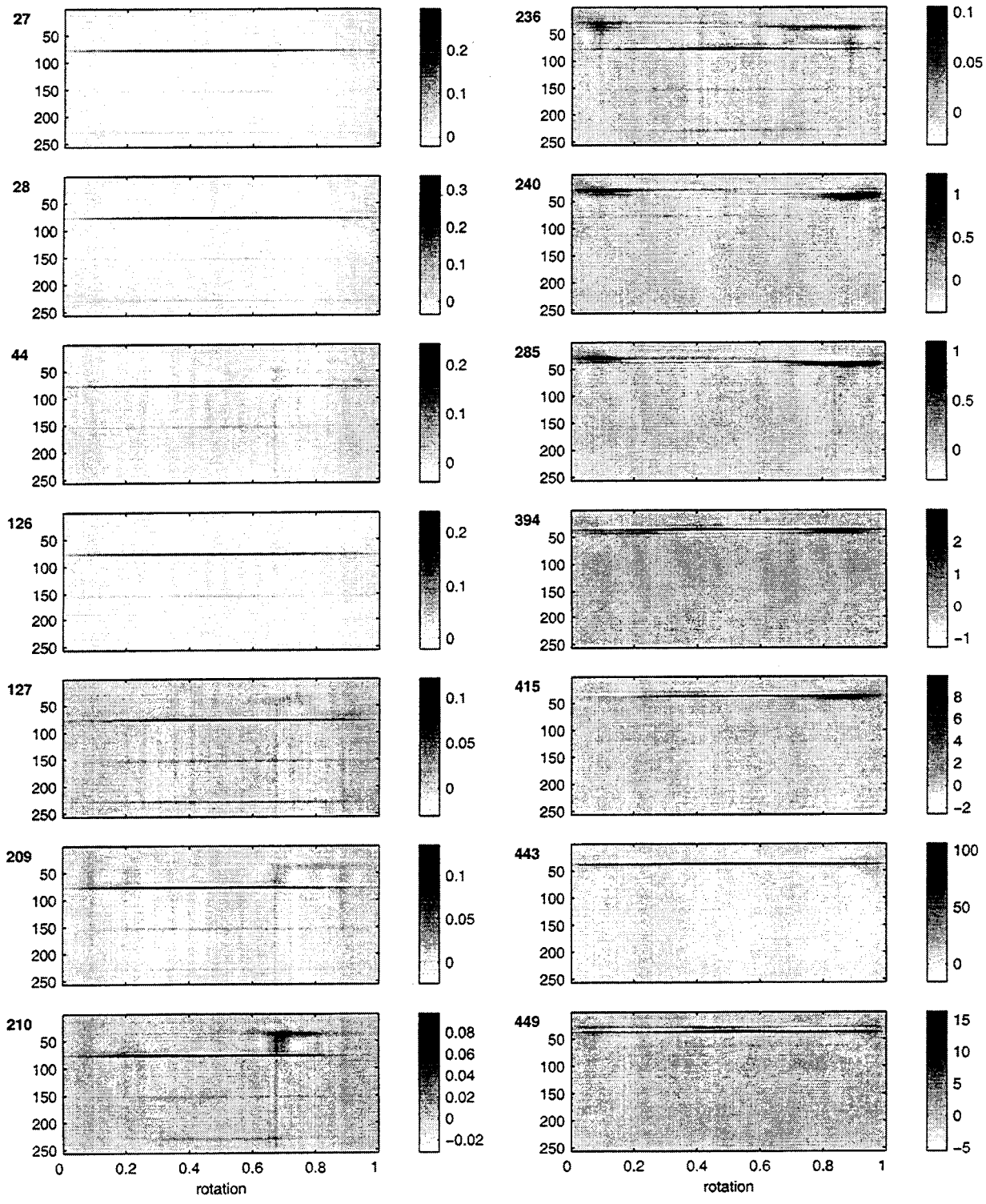


Figure A7. Wigner-Ville Distribution with Choi Williams Kernel of time synchronous average for one rotation of gear: Data from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

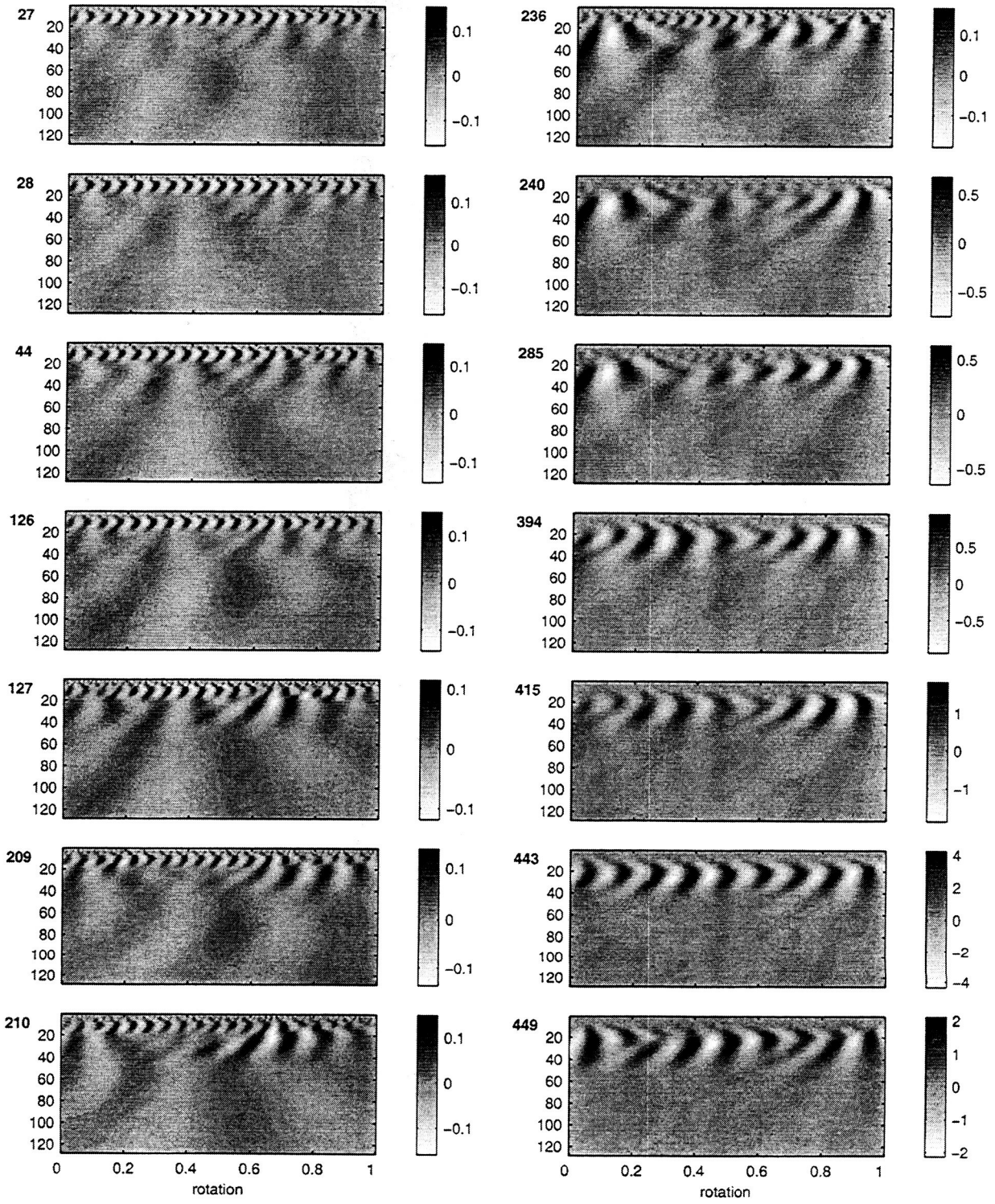


Figure A8. Continuous Wavelet Transform using Daubechies seventh order wavelet of time synchronous average for one rotation of gear: Data from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

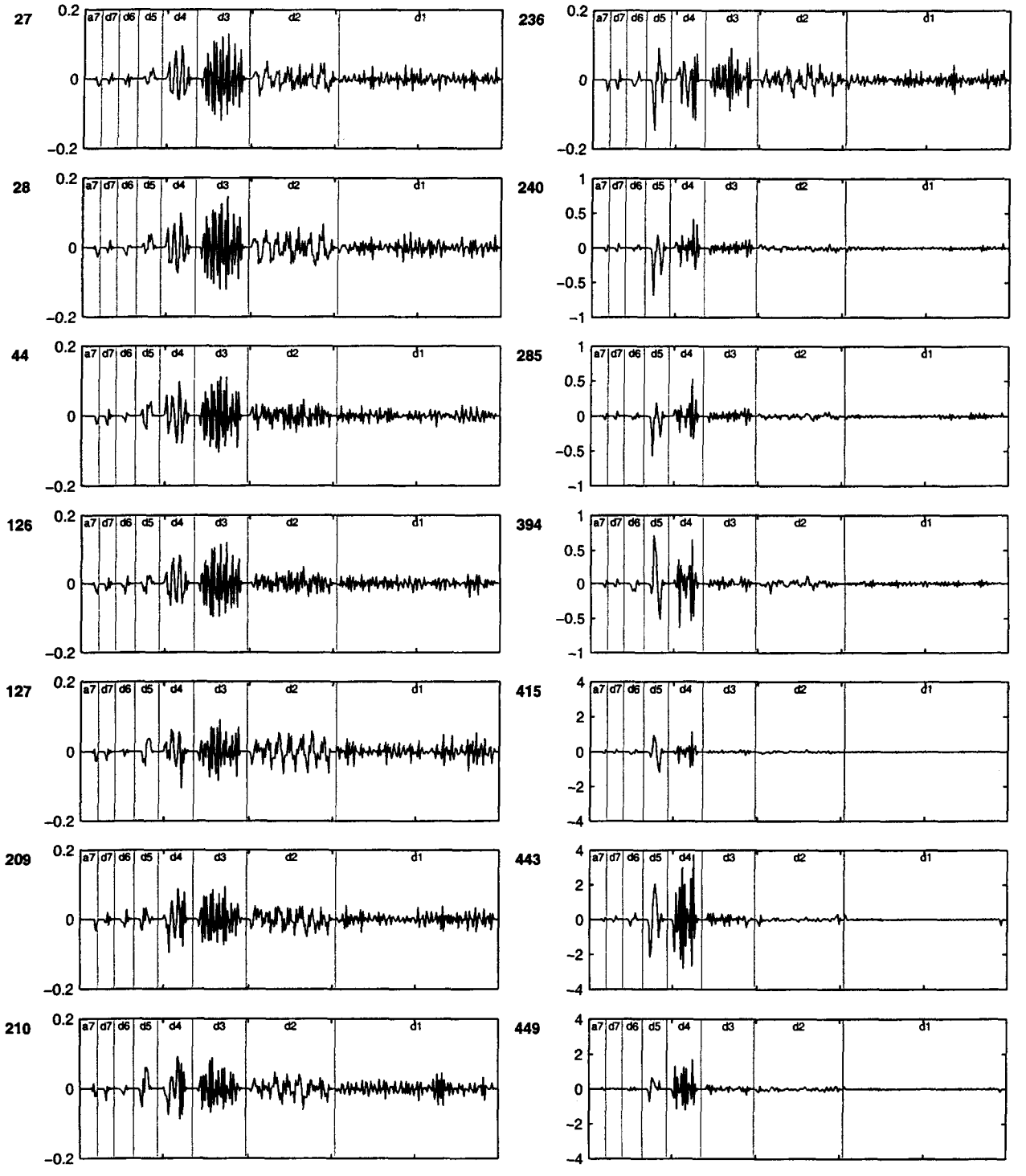


Figure A9. Discrete Wavelet Transform using Daubechies seventh order wavelet of time synchronous average for one rotation of gear: Data from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

APPENDIX B. RESIDUAL FROM SVD MODEL

Table B1. Residual from SVD Modeling of data from accelerometer measuring vibration of pinion gear.

record #	Time	Spectrum	STFT	WV-CW	CWT	DWT
27	0.20148	0.11079	0.19351	0.054207	0.12368	0.20148
28	0.22727	0.10897	0.23353	0.070008	0.38918	0.22727
44	0.33562	0.1485	0.28923	0.10801	0.39448	0.33562
126	0.22498	0.11899	0.21426	0.055327	0.1444	0.22498
127	0.54136	0.42919	0.42571	0.49581	0.50781	0.54136
209	0.5459	0.38619	0.48313	0.32999	0.61374	0.5459
210	0.72808	0.52639	0.68689	0.60371	0.71191	0.72808
236	0.78137	0.61549	0.75569	0.73746	0.79978	0.78137
240	0.97043	0.89005	0.96247	0.99429	0.96665	0.97043
285	0.95984	0.88667	0.9555	0.99325	0.94717	0.95984
394	0.99019	0.93515	0.99097	0.99886	0.98672	0.99019
415	0.99609	0.95112	0.99652	0.9992	0.99522	0.99609
443	0.99729	0.98408	0.99607	0.99973	0.99125	0.99729
449	0.99819	0.96179	0.99508	0.99912	0.99911	0.99819

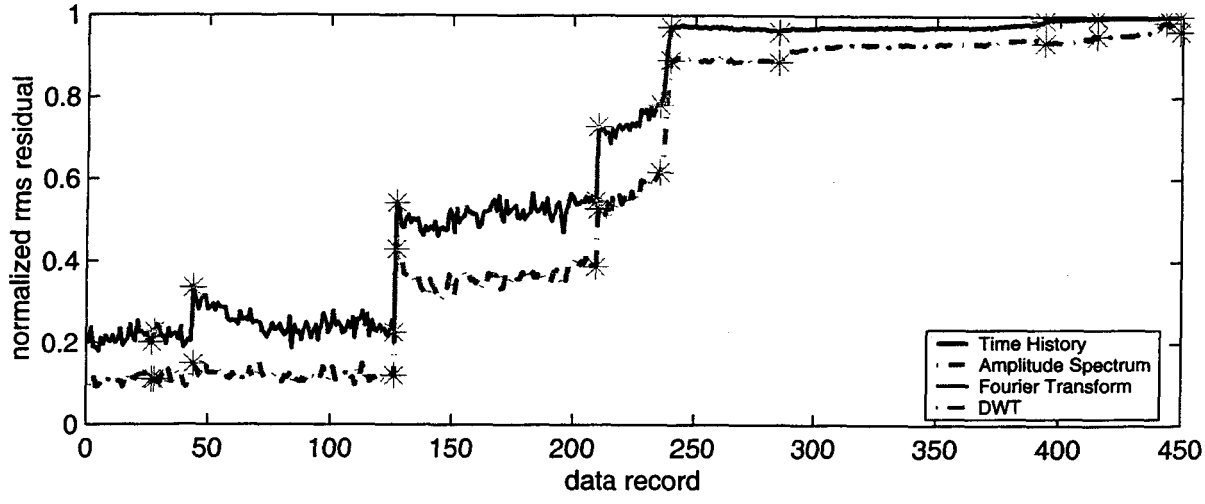


Figure B1. Normalized rms residual from SVD models of Time History, Amplitude Spectra, Fourier Transform and Discrete Wavelet Transform.

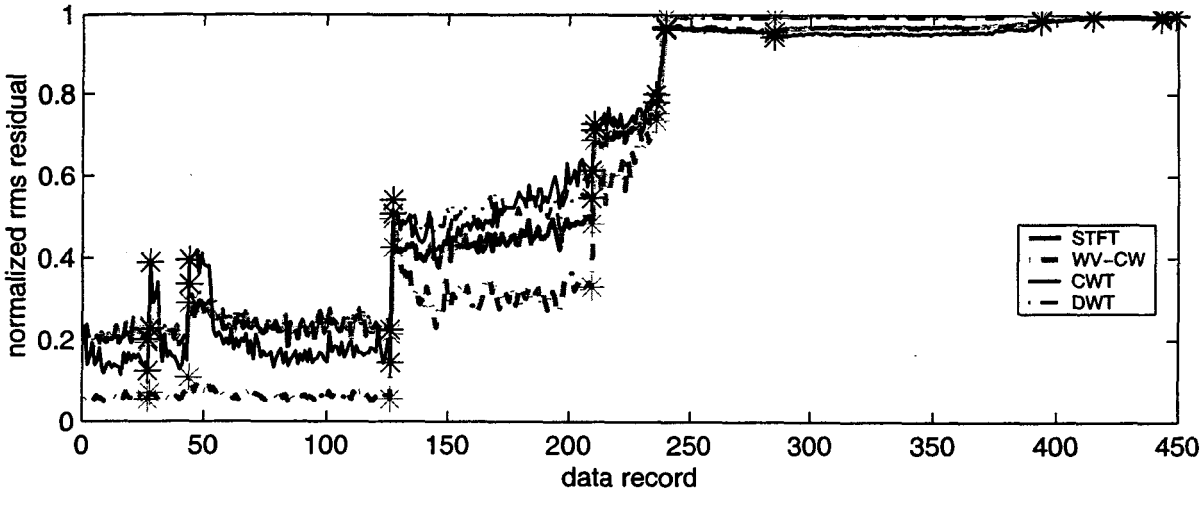


Figure B2. Normalized rms residual from SVD models of Short Time Fourier Transform, Wigner-Ville distribution with Choi-Williams kernel, Continuous Wavelet Transform and Discrete Wavelet Transform.

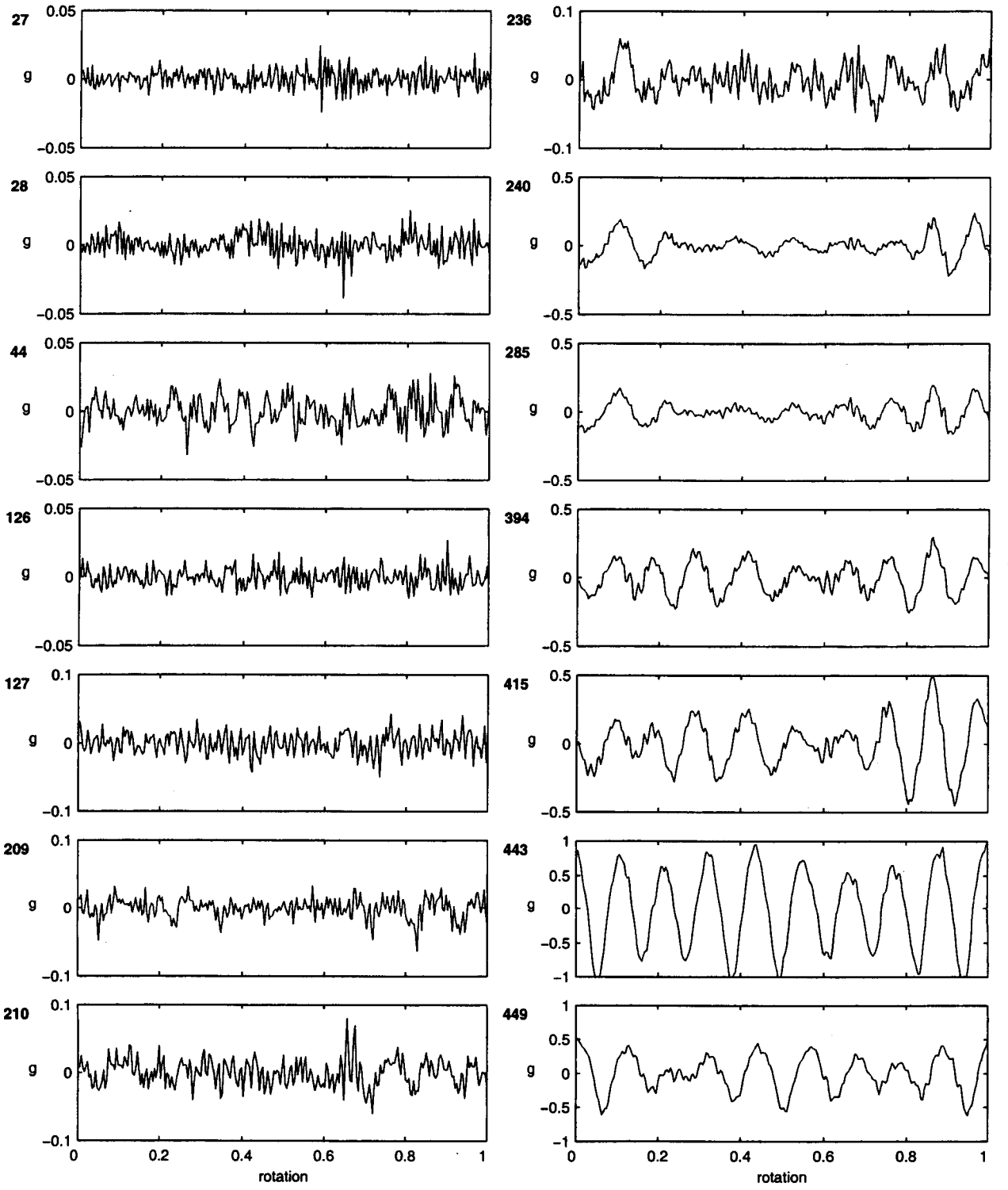


Figure B3. Time synchronous average for one rotation of gear: Residual from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

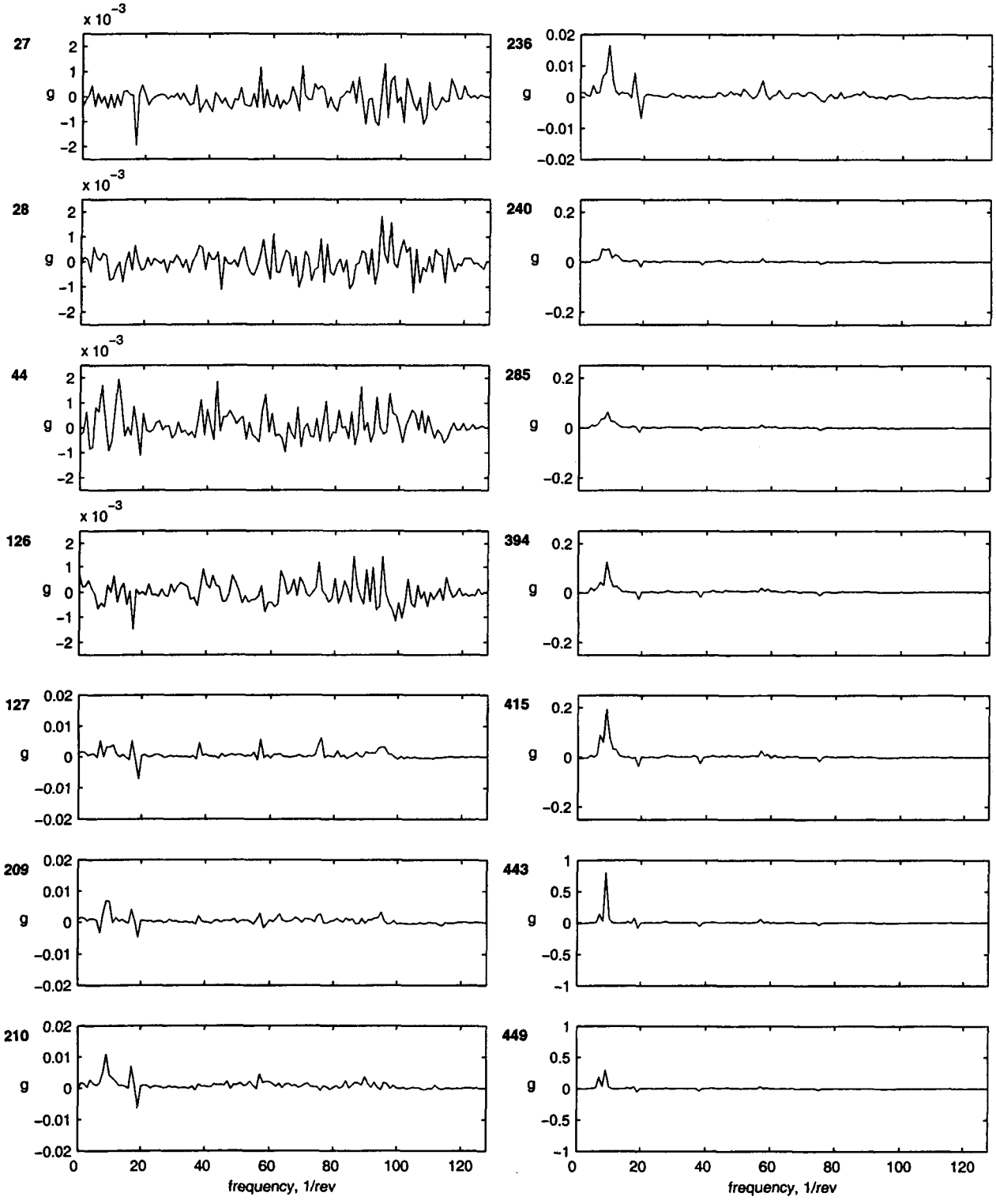


Figure B4. Amplitude of Fourier Transform of time synchronous average for one rotation of gear: Residual from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

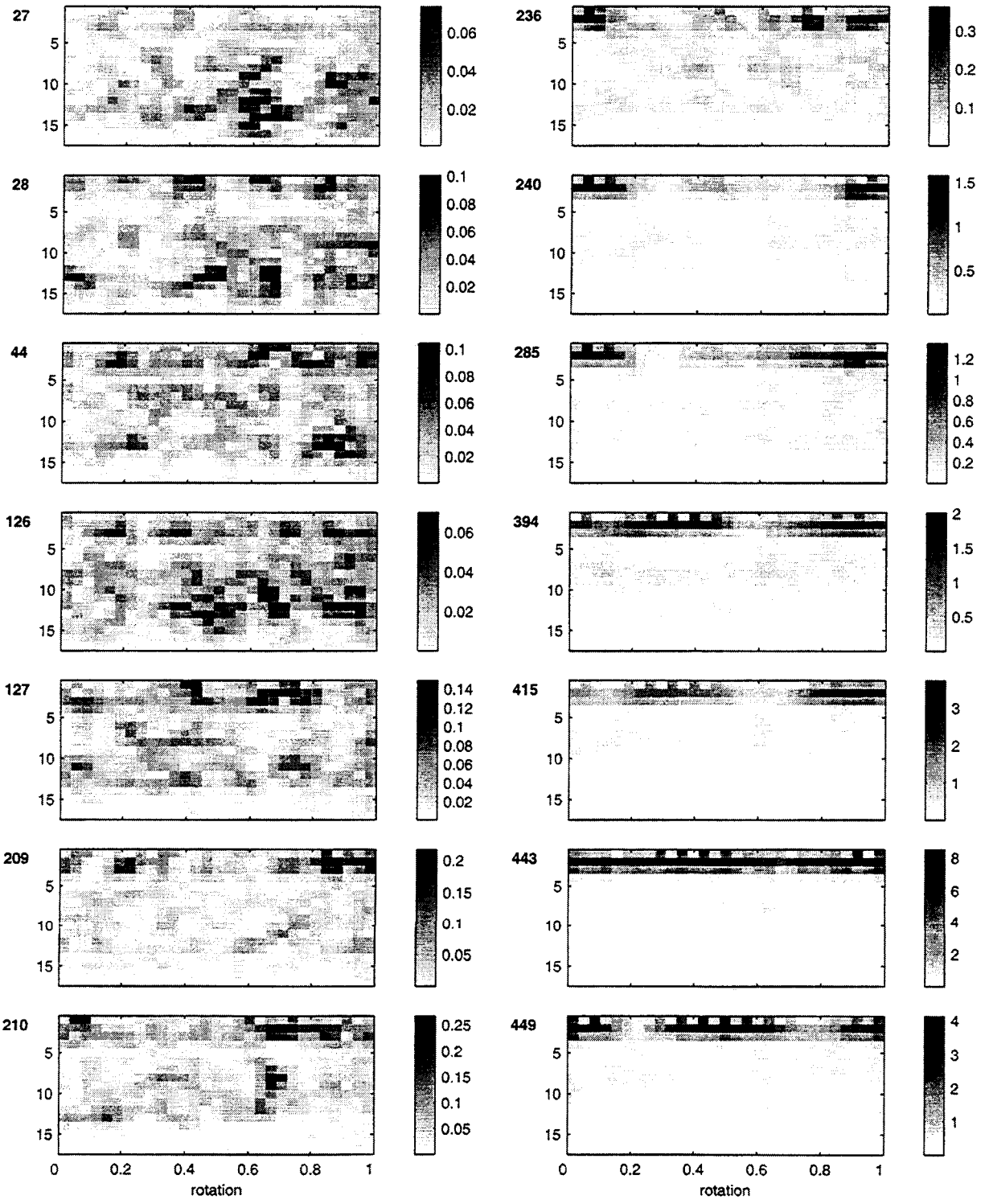


Figure B5. Short Time Fourier Transform of time synchronous average for one rotation of gear: Residual from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

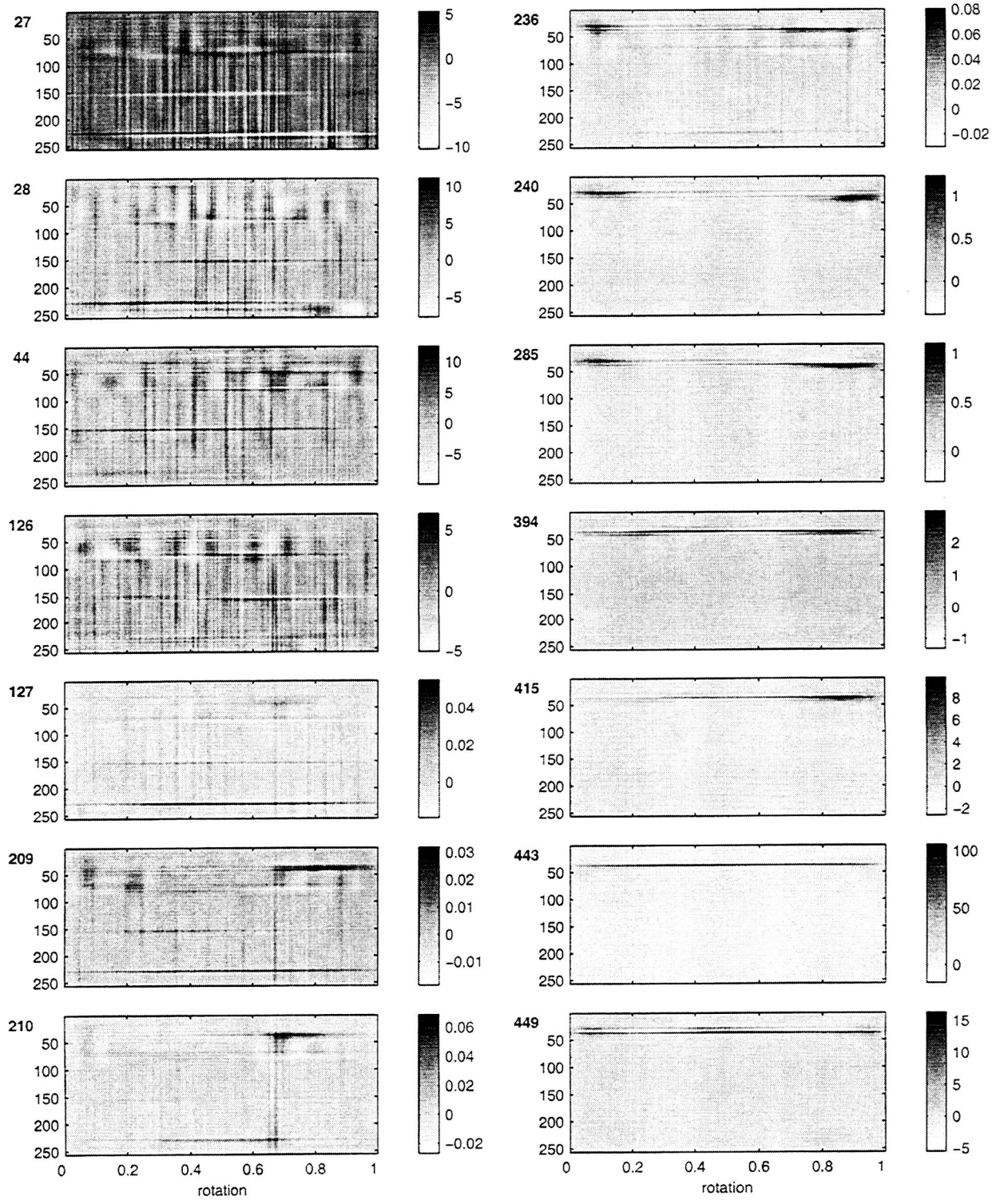


Figure B6. Wigner-Ville Distribution with Choi Williams Kernel of time synchronous average for one rotation of gear: Residual from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

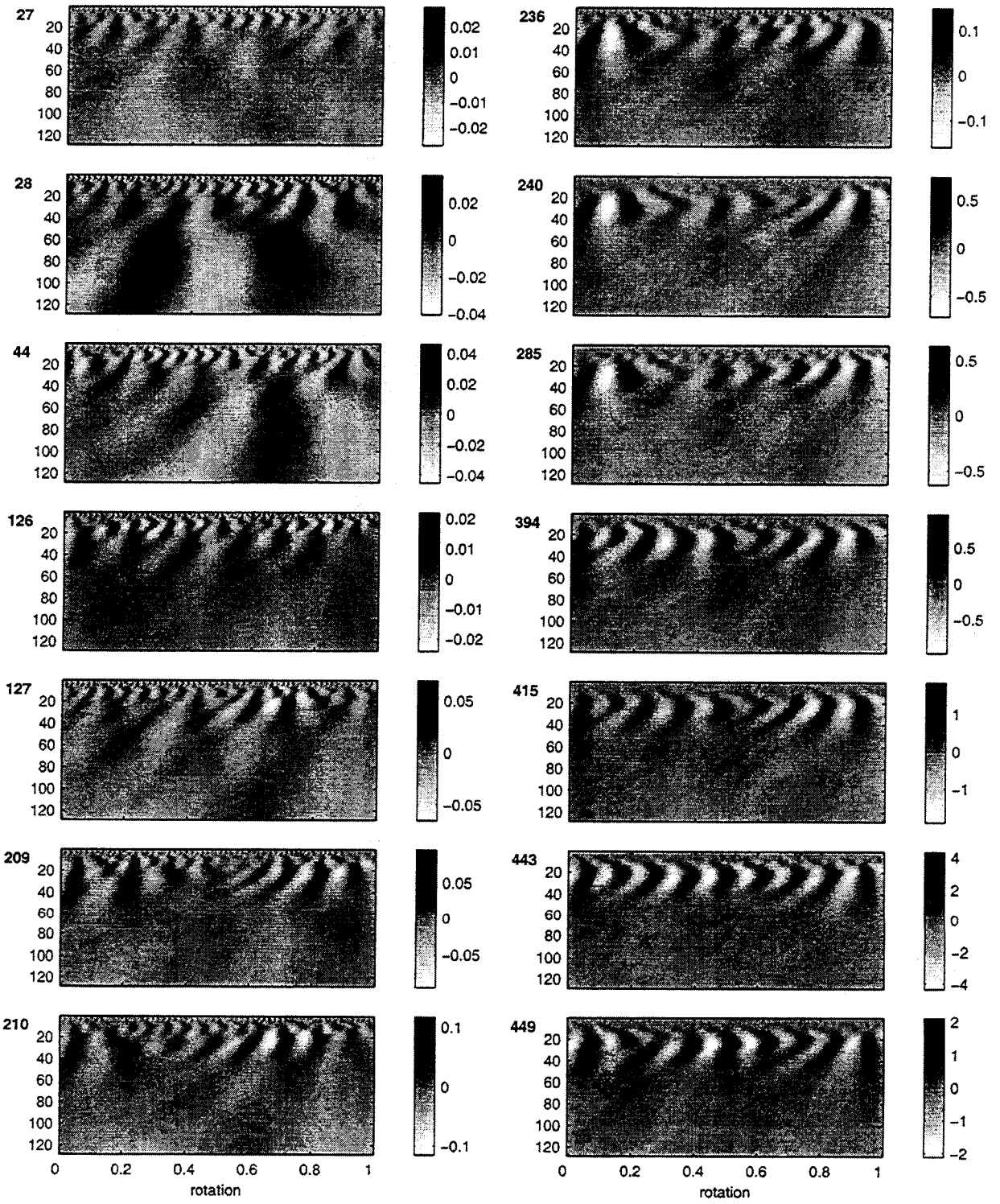


Figure B7. Continuous Wavelet Transform using Daubechies seventh order wavelet of time synchronous average for one rotation of gear: Residual from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

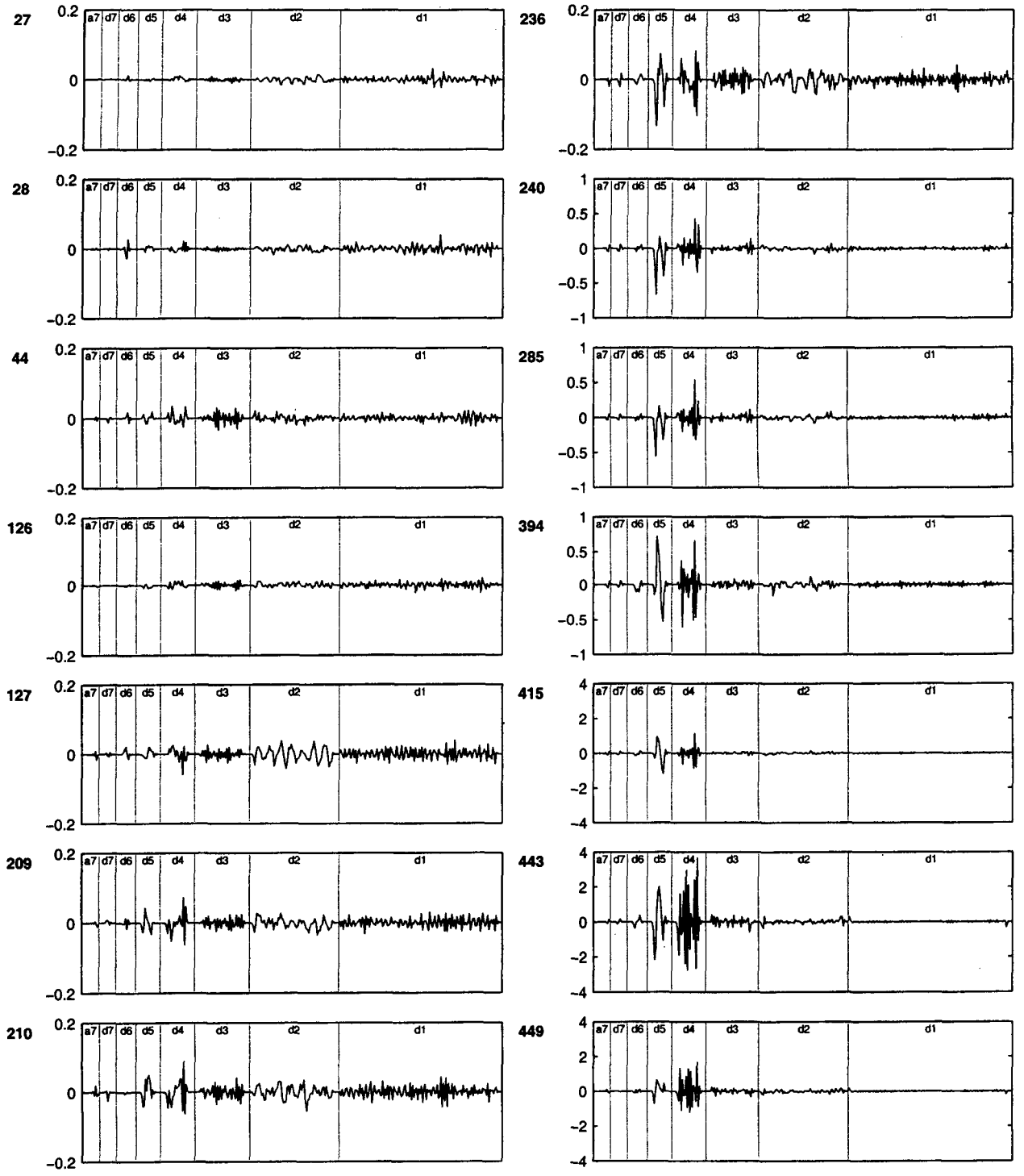


Figure B8. Discrete Wavelet Transform using Daubechies seventh order wavelet of time synchronous average for one rotation of gear: Residual from record numbers 27, 28, 44, 126, 127, 209, 210, 236, 240, 285, 394, 415, 443, and 449.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2003	Technical Memorandum

4. TITLE AND SUBTITLE
Detailed Vibration Analysis of Pinion Gear with Time-Frequency Methods

5. FUNDING NUMBERS
UPN 704-01-32

6. AUTHOR(S)
Marianne Mosher, Anna H. Pryor and David G. Lewicki

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Ames Research Center
Moffett Field, CA 94035-1000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
NASA/TM–2003-212269

11. SUPPLEMENTARY NOTES
Point of Contact: Marianne Mosher, Ames Research Center, MS 269-1, Moffett Field, CA 94035-1000
(650) 604-4055

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified — Unlimited
Subject Category 59 Distribution: Standard
Availability: NASA CASI (301) 621-0390

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 words)***

In this paper, the authors show a detailed analysis of the vibration signal from the destructive testing of a spiral bevel gear and pinion pair containing seeded faults. The vibration signal is analyzed in the time domain, frequency domain and with four time-frequency transforms: the Short Time Frequency Transform (STFT), the Wigner-Ville Distribution with the Choi-Williams kernel (WV-CW), the Continuous Wavelet Transform (CWT) and the Discrete Wavelet Transform (DWT). Vibration data of bevel gear tooth fatigue cracks, under a variety of operating load levels and damage conditions, are analyzed using these methods. A new metric for automatic anomaly detection is developed and can be produced from any systematic numerical representation of the vibration signals. This new metric reveals indications of gear damage with all of the time-frequency transforms, as well as time and frequency representations, on this data set. Analysis with the CWT detects changes in the signal at low torque levels not found with the other transforms. The WV-CW and CWT use considerably more resources than the STFT and the DWT. More testing of the new metric is needed to determine its value for automatic anomaly detection and to develop fault detection methods for the metric.

14. SUBJECT TERMS
vibration analysis, anomaly detection, time-frequency methods, rotorcraft, pinion gear, gear damage detection

15. NUMBER OF PAGES
47

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified		